PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/61, 15/62, 9/90, A61K 7/48, 7/06

A1

(11) International Publication Number: **WO 00/70064**

(43) International Publication Date: 23 November 2000 (23.11.00)

(21) International Application Number: PCT/DK00/00265

(22) International Filing Date: 17 May 2000 (17.05.00)

(30) Priority Data:

PA 1999 00683	17 May 1999 (17.05.99)	DK
PA 1999 00689	18 May 1999 (18.05.99)	DK
60/137,068	2 June 1999 (02.06.99)	US

(71) Applicant *(for all designated States except US)*: NOVO NORDISK A/S [DK/DK]; Enzyme Business Patents, Novo Allé, DK–2880 Bagsværd (DK).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: HJORT, Carsten, Mailand [DK/DK]; Råbroparken 36, DK–2765 Smørum (DK).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: POLYPEPTIDES WITH PROTEIN DISULFIDE REDUCING PROPERTIES

(57) Abstract

There is provided a protein disulfide isomerase variant having increased reducing properties as compared to the wild–type protein and capable of being expressed extracellularly in a high yield. The present invention relates to polypeptides capable of reducing disulfide bonds, including variants of protein disulfide isomerase, the variants having a decreased redox potential. The invention also relates to nucleotide sequences encoding the polypeptides, as well as nucleic acid constructs, vectors, and host cells comprising the nucleotide sequences. Further, the invention relates to methods for producing and using the polypeptides, including use of the polypeptides for reducing the allergenicity of allergenic proteins, e.g. in feed or food or cosmetic products. The invention also relates to food additives or cosmetic products comprising a polypeptide of the invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

POLYPEPTIDES WITH PROTEIN DISULFIDE REDUCING PROPERTIES

FIELD OF THE INVENTION

The present invention relates to polypeptides capable of reducing disulfide bonds, including variants of protein disulfide isomerase. The invention also relates to nucleotide sequences encoding the polypeptides, as well as nucleic acid constructs, vectors, and host cells comprising the nucleotide sequences. Further, the invention relates to methods for producing and using the polypeptides, including use of the polypeptides for reducing the allergenicity of allergenic proteins.

BACKGROUND OF THE INVENTION

Many of the proteins responsible for allergic reactions have intramolecular disulfide bonds. Disulfide bonds in proteins are formed between cysteine residues and have the function of stabilising the secondary and tertiary structure of the protein. By reducing these bonds it has been shown that it is possible to change the allergic properties of a given protein.

Among the protein disulfide redox enzymes are two main groups of enzymes, thioredoxin (TRX) and protein disulfide isomerase (PDI). The redox enzymes catalyse the general reaction:

$$R_1\text{-SH} + R_2\text{-SH} + Enz_{ox} \leftrightarrow R_1\text{-S-S-}R_2 + Enz_{red}$$

where R_1 and R_2 represent protein entities which are the same or different, either within the same polypeptide or in two polypeptides, Enz_{ox} is a protein disulfide redox enzyme in the oxidised state and Enz_{red} is a protein disulfide redox enzyme in the reduced state.

EC 5.3.4.1 (Enzyme Nomenclature, Academic Press, Inc. 1992), Protein disulfide isomerase (alternative named S-S rearrangase) is an enzyme capable of catalysing the rearrangement of both intrachain and interchain –S-S- bonds in proteins. In order words, protein disulfide isomerase catalyzes the oxidation, reduction and isomerization of protein disulfides. Most of the enzymes capable of reducing disulfide bonds have the following amino acid sequence in common:

$$R^1 - Cys - X_1 - Y_1 - Cys - R^2 \quad \text{(Formula 1)}$$

wherein R^1 and R^2 each are different amino acid sequences. For protein disulfide isomerase (PDI) generally X_1 is Gly, and Y_1 is His, and for thioredoxin (TRX) gener-

ally X_1 is Gly, and Y_1 is Pro, using the conventional nomenclature for amino acid residues. In nature, PDI consists of two subunits, each consisting of two homologous domains, where each domain comprises Formula 1 depicted above, whereas TRX consists of one domain only. The crystal structure of a protein disulfide isomerase has been reported in "Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*, McCarthy A. A. et al., 2000, Nat Struct Biol, 7(3):196-9".

It has been shown that mitigation of food allergy is achieved by the reduction of disulfide bonds using thioredoxin (TRX), e.g. US 5,792,506 or B.B. Buchanan et al. (Proc. Natl. Acad. Sci, USA, Vol. 94, pp. 5372-5377, 1997) and WO 96/12799 describing experiments performed on sensitised dogs fed with TRX treated food, showing TRX as a potent reductant of allergenic proteins. However, the use of thioredoxin on an industrial scale for the production of low allergy food remains to be exploited. This is due to the fact that it has not yet been possible to achieve extracellularly expressed disulfide reducing proteins in quantities large enough for industrial application.

In nature, protein disulfide isomerase is expressed intracellularly only, and has proven very difficult to express extracellularly due to proteolytic degradation.

Extracellularly expressed variants have been described previously, for example in WO 95/00636 disclosing a fungal protein disulfide isomerase from *Aspergillus* and sequences for the recombinant production of the protein, and in WO 95/01425 describing compositions comprising protein disulfide redox enzymes, both applications describing the use of the protein variants for the treatment or degradation of in particular scleroproteins, especially in hair, skin and wool.

There is a need for providing the combination of a highly disulfide reducing enzyme capable of large-scale recombinant production with a high output, which has so far not been possible.

SUMMARY OF THE INVENTION

The inventor has provided a protein disulfide isomerase variant having increased reducing properties as compared to the wild-type protein and capable of being expressed extracellularly in a high yield.

Accordingly, the invention relates to polypeptides capable of reducing disulfide bonds in proteins, which polypeptide comprises or consists of an amino acid se-

quence having (i) at least 60% similarity with the amino acid sequence set forth in amino acid number 21-281 of SEQ ID NO:15; or (ii) at least 60% similarity with the amino acid sequence set forth in SEQ ID NO:17; provided that a position in the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is selected from the group consisting of: i) Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys, iii) Cys-Val-Leu-Cys, and iv) Cys-Gly-Tyr-Cys.

The invention further relates to a polypeptide capable of reducing disulfide bonds which polypeptide is a variant of a parent protein disulfide isomerase by having an amino acid sequence differing from that of the parent protein disulfide isomerase in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13, which position is Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys, iii) Cys-Val-Leu-Cys, or iv) Cys-Gly-Tyr-Cys for the variant polypeptide.

The present invention also relates to a method for providing a polypeptide capable of reducing disulfide bond and to the polypeptides obtainable by such methods, which polypeptide is a variant of a protein disulfide isomerase, the method comprising the step of: altering at least one amino acid residue by amending the amino acid X_1 to Glu and/or the amino acid Y_1 to Pro in the active site corresponding to -Cys-X_1-Y_1-Cys- of a parent protein disulfide isomerase; to obtain a variant of the protein disulfide isomerase comprising Cys-Gly-Pro-Cys.

The invention further relates to a polypeptide, which is a protein disulfide isomerase variant having the formula of: $X - A - [Z - B]_n - Y$,
wherein (i) "A" is the amino acid sequence: Cys-Gly-Pro-Cys; (ii) "X" is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:1 or a functional equivalent thereof; (iii) "Y" is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:2 or a functional equivalent thereof; (iv) "Z" is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:11 or a functional equivalent thereof; (v) "B" is individually selected from the amino acid sequence Cys-Gly-Pro-Cys or Cys-Gly-His-Gly or Cys-Ala-Thr-Cys or Cys-Pro-His-Cys or Cys-Val-Leu-Cys or Cys-Gly-Tyr-Cys; and (vi) n = 0 or an integer from 1 to 3.

The present invention also relates to nucleotide sequences encoding the polypeptides of the invention and to nucleic acid constructs, vectors, and host cells comprising the nucleotide sequences as well as methods for producing the polypeptides.

Furthermore, the invention relates to compositions comprising the polypeptides of the invention, and to the use of the polypeptides, e.g. for reducing the allergenicity of an allergenic protein, e.g. in food and feed products and cosmetics, or for increasing the digestibility of food or feed. The invention further relates to food or feed additives comprising a polypeptide of the invention.

DRAWINGS

Figure 1 shows the construction of the plasmid pCaHj 527.

Figure 2 depicts the construction of the plasmid pCaHj 548.

Figure 3 shows the structure of A. oryzae PDI.

Figure 4 shows alignment of amino acid sequences of a PDI from *Aspergillus oryzae* with a PDI from *T. reseii*, *H. insolens*, *S. cerevisiae* and *A. niger*. Percent Similarity in upper triangle. Percent Divergence in lower triangle.

Figure 5 shows results from Examples, Comparison of recombinant fungal protein disulfide isomerase (PDI) and DsbA isomerase.

Figure 6 shows results from Examples, Optimisation of the 'PDI:substrate' ratio.

Figure 7 shows alignment of 7 PDI gene product, cf. Examples.

DETAILED DESCRIPTION OF THE INVENTION

Polypeptides of the invention and methods of providing such polypeptides

The inventor has provided a protein disulfide isomerase variant (SEQ ID NO:15, where amino acid number 21-281 of SEQ ID NO:15 is the polypeptide without the signal peptide) having increased capacity for reducing protein disulfide bond and capable of being expressed extracellularly in a high yield.

Accordingly, within the scope of the invention is a polypeptide capable of reducing disulfide bonds, the polypeptide selected from the group of:

(a) a polypeptide consisting of or comprising an amino acid sequence having at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence of SEQ ID

NO:15 or the amino acid sequence set forth in amino acid number 21-281 of SEQ ID NO:15;

(b) a polypeptide consisting of or comprising an amino acid sequence having at least 60% at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence set forth in SEQ ID NO:17;

(c) a polypeptide capable of reducing disulfide bonds, which polypeptide comprises or consists of an amino acid sequence encoded by a nucleotide sequence which hybridizes under low, medium or high stringency conditions with (i) SEQ ID NO:14; (ii) SEQ ID NO:16; (iii) a subsequence of (i) or (ii) of at least 100 nucleotides; or (iv) a complementary strand of (i), (ii) or (iii); and

(d) a polypeptide capable of reducing disulfide bonds which (i) is encoded by the nucleotide sequence contained in plasmid pCaHj548; or (ii) comprises or consist of an amino acid sequence having at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence encoded by the nucleotide sequence contained in plasmid pCaHj548;

provided that a position of the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO: 13 is selected from the group consisting of: i) Cys-Gly-Pro-Cys; ii) Cys-Ala-Thr-Cys; iii) Cys-Val-Leu-Cys; and iv) Cys-Gly-Tyr-Cys.

SEQ ID NO:12 is the nucleotide sequence of a wt. *A. oryzae* PDI gene where the coding sequence is nucleotide number 71-444, 502-880, 692-1401 and 1478-1832 of SEQ ID NO:12. The corresponding wt-PDI amino acid sequence is shown in SEQ ID NO:13, where amino acid number 1-20 is the signal peptide. The active sites of the wt. *A. oryzae* PDI (i.e. i.a. amino acid number 58-61 of SEQ ID NO:13) is coded by nucleotides 242-253 and 1385-1396 of SEQ ID NO:12.

The polypeptide SEQ ID NO:15 is identical to amino acid number 1 to 281 of SEQ ID NO:13, except that the polypeptide given by SEQ ID NO:15 has the sequence Cys-Gly-Pro-Cys in the position corresponding to 58-61 of SEQ ID NO: 13, i.e. amino acid number 61 is Pro in SEQ ID NO:15 and amino acid number 61 is His

in SEQ ID NO:13. SEQ ID NO:14 is the nucleotide sequence coding for the amino acid sequence given by SEQ ID NO:15. The amino acid number 1-20 of SEQ ID NO:15 is a signal peptide. Amino acid number 21-281 of SEQ ID NO:15 is the mature variant PDI of the invention, i.e. without the signal peptide. The term " SEQ ID NO:15" as used herein may also denote the amino acid sequence of SEQ ID NO:15 without the signal sequence, i.e. the amino acid sequence given by amino acid number 21-281 of SEQ ID NO:15.

The polypeptide SEQ ID NO:17 is identical to amino acid number 21 to 115 of SEQ ID NO:13, except that the polypeptide given by SEQ ID NO:17 has Cys-Gly-Pro-Cys in the position corresponding to 58-61 in SEQ ID NO: 13. SEQ ID NO:16 is the nucleotide sequence coding for the amino acid sequence SEQ ID NO:17.

The amino acid sequence identity is the degree of identity between the two sequences indicating a derivation of the first sequence from the second. The amino acid sequence similarity also takes into account conservative amino acid substitutions. The degree of identity or similarity may suitably be determined by means of computer programs known in the art. Both the degree of identity and similarity takes into account possible gaps.

For purposes of the present invention, the degree of identity or similarity between two amino acid sequences may be determined by the Clustal method (Higgins, 1989, *CABIOS* 5: 151-153) using the LASERGENE™ MEGALIGN™ software (DNASTAR, Inc., Madison, WI) with an identity table and the following multiple alignment parameters: Gap penalty of 10 and gap length penalty of 10. Pairwise alignment parameters were Ktuple=1, gap penalty=3, windows=5, and diagonals=5].

The degree of identity or similarity may also be determined according to the method described in Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-45, with the following settings for amino acid sequence comparison: GAP creation penalty of 3.0 and GAP extension penalty of 0.1. The determination may be done by means of a computer program known such as GAP provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711). Two given sequences can be aligned according to the method de-

scribed in Needleman (*supra*) using the same parameters. This may be done by means of the GAP program (*supra*).

The degree of hybridisation may be determined by the method described in J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York.

For long probes of at least 100 nucleotides in length, very low to very high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures.

For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2 x SSC, 0.2% SDS preferably at least at 45°C (very low stringency), more preferably at least at 50°C (low stringency), more preferably at least at 55°C (medium stringency), more preferably at least at 60°C (medium-high stringency), even more preferably at least at 65°C (high stringency), and most preferably at least at 70°C (very high stringency).

For short probes which are about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at about 5°C to about 10°C below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures.

For short probes, which are about 15 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6X SCC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6X SSC at 5°C to 10°C below the calculated T_m.

In one preferred embodiment, the polypeptide of the invention has Cys-Gly-Pro-Cys in the position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

The polypeptide of the invention may comprise SEQ ID NO:15, the amino acid sequence of amino acid number 21-281 of SEQ ID NO:15, or SEQ ID NO:17. In a further embodiment, the polypeptide of the invention consists of the amino acid sequence selected from the group consisting of: (i) SEQ ID NO:13 with Cys-Gly-Pro-Cys in amino acid residues numbered 58-61; (ii) SEQ ID NO:13 without amino acid number 1-20 and with Cys-Gly-Pro-Cys in amino acid residues numbered 58-61 in SEQ ID NO:13; (iii) SEQ ID NO:15; (iv) amino acid number 21-281 of SEQ ID NO:15; and (v) SEQ ID NO:17.

The polypeptide of the invention may also be a protein disulfide isomerase having an amino acid sequence which in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys; iii) Cys-Val-Leu-Cys; or iv) Cys-Gly-Tyr-Cys. In a preferred embodiment the polypeptide is a protein disulfide isomerase having an amino acid sequence, which in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is Cys-Gly-Pro-Cys.

The invention also relates to a polypeptide capable of reducing disulfide bonds, which is a variant of the polypeptides as described herein, such as a variant, e.g., of SEQ ID NO:15; a variant of the amino acid sequence set forh in amino acid number 21-281 of SEQ ID NO:15; or a variant of SEQ ID NO:17, the variant comprising substitution(s), deletion(s), and/or insertion(s) of one or more amino acids; provided that the variant comprises an amino acid sequence selected from the group consisting of: i) Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys, iii) Cys-Val-Leu-Cys, and iv) Cys-Gly-Tyr-Cys, in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

The polypeptides of the invention may, e.g., further comprise an amino acid sequence which has at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identity or similarity with the amino acid sequence set forth in SEQ ID NO:18 (i.e. amino acid number 363 to 464 of SEQ ID NO:13) provided the position corresponding to amino acid residues numbered 31-34 in SEQ ID NO:18 is selected from the group consisting of: (i) Cys-Gly-Pro-Cys, (ii) Cys-Ala-Thr-Cys, (iii) Cys-Val-Leu-Cys, (iv) Cys-Gly-Tyr-Cys, and (v) Cys-Gly-His-Cys.

Thus, the polypeptide of the invention may, e.g., have one, two or more domains, e.g. tree domains, each domain comprising a sequence of the type Cys-X_1-Y_1-Cys, wherein the polypeptide comprises at least one Cys-X_1-Y_1-Cys selected from the group consisting of: i) Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys, iii) Cys-Val-Leu-Cys, and iv) Cys-Gly-Tyr-Cys. In one embodiment the polypeptide has two "Cys-X_1-Y_1-Cys" regions, both of which are Cys-Gly-Pro-Cys. In other embodiments, the polypeptides of the invention have only *one* domain comprising the region "Cys-X_1-Y_1-Cys", which region is Cys-Gly-Pro-Cys.

It is understood that Cys-X_1-Y_1-Cys is in a position of the polypeptide conferring catalytic activity, i.e. in a position corresponding to the amino acid residues numbered 58-61 in SEQ ID NO:13, which position can be determined by the person skilled in the art by alignment of the amino acid sequence in question with SEQ ID NO:13.

The polypeptide of the invention may be a protein disulfide isomerase (PDI, EC 5.3.4.1.).

The invention also provides a polypeptide capable of reducing disulfide bonds, the polypeptide being a variant of a parent protein disulfide isomerase by having an amino acid sequence differing from that of the parent protein disulfide isomerase in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 (i.e. Cys-X_1-Y_1-Cys-), which position for the variant polypeptide is selected from the group consisting of i) Cys-Gly-Pro-Cys, ii) Cys-Ala-Thr-Cys, iii) Cys-Val-Leu-Cys, and iv) Cys-Gly-Tyr-Cys.

The amino acid sequence of the polypeptide of the invention may - apart from a difference in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 - also differ from the amino acid sequence of the parent protein disulfide isomerase by having further substitution(s), deletion(s), and/or insertion(s) of one or more amino acids compared to the parent protein disulfide isomerase, provided that the polypeptide comprises the amino acid sequence Cys-Gly-Pro-Cys, Cys-Ala-Thr-Cys, Cys-Val-Leu-Cys, or Cys-Gly-Tyr-Cys in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13. The amino acid sequence of the polypeptide being a variant of a parent protein disulfide isomerase may have at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity or similarity with

the amino acid sequence of the parent protein disulfide isomerase. Also contemplated are methods for providing such polypeptides which are variants of a parent protein disulfide isomerase (PDI) by amending the amino acid sequence of the parent PDI as described herein.

The parent protein disulfide isomerase may be any PDI enzyme, i.e. an enzyme that is classified as a PDI under EC 5.3.4.1. The parent PDI may be a fungal PDI, such as, e.g. a filamentous fungal PDI, e.g. native to a strain of *Humicola, Fusarium or Aspergillus*, such e.g. *H. insolens, F. solani pisi* or *A. oryzae*. The parent PDI, may be derived from *T. reseii*, *H. insolens*: *S. cerevisiae*: or A. niger, the sequences of which in Figure 4 are compared by alignment to an *A. oryzae* PDI (SEQ ID NO:13). The parent PDI may be SEQ ID NO:13, a subsequence of SEQ ID NO:13 having protein disulfide isomerase activity, such as, e.g. amino acid no. 21-115 of SEQ ID NO:13, e.g. amino acid no. 1-115 of SEQ ID NO:13, e.g. amino acids no. 1-281 of SEQ ID NO:13, e.g. amino acids no. 21-281 of SEQ ID NO:13, or a subsequence thereof having protein disulfide isomerase activity. The parent protein disulfide isomerase may, e.g. be encoded by a nucleotide sequence which hybridises under low, medium or high stringency conditions with one or more of (i) SEQ ID NO:12, SEQ ID NO:14 or SEQ ID NO:16; (ii) nucleotides 131 to 365 of SEQ ID NO:12; (iii) a subsequence of (i) or (ii) of at least 100 nucleotides; or (iv) a complementary strand of (i), (ii) or (iii).

Accordingly, within the scope of the invention is a method for providing a polypeptide capable of reducing disulfide bond, the method comprising the step of:

(a) altering at least one amino acid residue by amending the amino acid X_1 to Glu, Ala, Val, or Gly and/or by amending the amino acid Y_1 to Pro, Thr, Leu, or Tyr in the active site corresponding to Cys-X_1-Y_1-Cys of a parent protein disulfide isomerase;

to obtain a variant of the parent protein disulfide isomerase comprising Cys-Gly-Pro-Cys, Cys-Ala-Thr-Cys, Cys-Val-Leu-Cys, or Cys-Gly-Tyr-Cys, in a position corresponding to amino acid number 58-61 of SEQ ID NO:13.

By the term "and/or" in the context of "altering at least one amino acid residue by amending the amino acid X_1 to Glu *and/or* the amino acid Y_1 to Pro in the active site corresponding to Cys-X_1-Y_1-Cys of a parent protein disulfide isomerase to obtain a variant of the parent protein disulfide isomerase comprising Cys-Gly-Pro-Cys in a position corresponding to amino acid number 58-61 of SEQ ID NO:13" is understood

that depending on the identity of X_1 and Y_1 in the parent PDI, it may be needed to *amend only one of X_1 and Y_1* (if either X_1= Glu or Y_1= Pro in the parent PDI) or it may be needed to *amend both X_1 and Y_1* in order to achieve X_1= Glu and Y_1= Pro (if X_1 is not Glu and Y_1 is not Pro in the parent PDI). The term "amend" includes any method which may be used to change the identity of the amino acids, such as, but not limited to, substitutions by site-directed mutagenesis of a corresponding DNA sequence, shuffling, or synthesising the polypeptide with the amino acid sequence of the variant by solid or liquid phase synthesis.

The method may further to step (a) (i.e. outside Cys-X_1-Y_1-Cys) comprise the step (b) of substitution(s), deletion(s), and/or insertion(s) of one or more amino acids compared to the parent protein disulfide isomerase. Within the scope of the invention are polypeptides capable of reducing disulfide bonds obtainable by such method. The variant polypeptide may have an amino acid sequence which has at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity or similarity with the amino acid sequence of the parent protein disulfide isomerase.

The present invention also provides a variant of a parent protein disulfide isomerase having increased disulfide reducing properties when compared to the parent protein disulfide isomerase (e.g. a wild-type protein), the variant having the formula: "$X - A - [Z - B]_n - Y$" as described above.

When n=0, the formula corresponds to X-A-Y, wherein A is the amino acid sequence: Cys-Gly-Pro-Cys, and X and Y are as defined above. The reducing capability of a protein variant of the formula X-A-Y is increased as compared to wild-type protein disulfide isomerase. The active site of the variant comprises A, but the activity is dependant on a larger part of the rest of the sequences, in particular a part of Y. Without being bound by theory this is believed to be due to the presentation of the active site, such as the tertiary structure of the variant.

The amino acid sequence of A (Cys-Gly-Pro-Cys) may also be expressed as C-G-P-C by the symbols of the one-letter amino acid code. Further to the formula "X" is an amino acid sequence positioned prior to the active site A, comprising a sequence corresponding to SEQ ID NO: 1, or a functional equivalent thereof. The signal se-

quence of X (amino acid 1-20) according to the invention may be replaced by any other signal sequence.

As described above naturally occurring DPI comprises domains with each an active site. For several applications it is preferred that the protein variant according to the invention comprises at least two active sites, corresponding to the formula above, wherein n=1. The active sites are separated by a sequence, whereby the positioning of the active sites is optimized with respect to presentation. The second active site, i.e. B, may be individually selected from the amino acid sequence Cys-Gly-Pro-Cys or Cys-Gly-His-Gly or Cys-Ala-Thr-Cys or Cys-Pro-His-Cys or Cys-Val-Leu-Cys or Cys-Gly-Tyr-Cys. The selection of the specific sequence depends on the modulation of the reducing properties of the protein variant as compared to the rearrangement properties.

Additional sequences B may be comprised in the protein variant, each separated from the other(s), as exemplified by the above formula, wherein n=2 and n=3. Any of the above-mentioned active sites may be present as identical repeats or they may be present as sequences different from each other, depending on their use.

The separating sequence, Z, is according to the invention an amino acid sequence positioned between A and B or B and B, i.e. active sites, comprising a sequence corresponding to SEQ ID NO: 11, or a functional equivalent thereof, wherein the functionality in particular relates to the presentation of the active site. Accordingly, in one embodiment of the [Z-B] alignment may be present in from 1 to 3 repeats, wherein B is individually selected from Cys-Gly-Pro-Cys or Cys-Gly-His-Gly or Cys-Ala-Thr-Cys or Cys-Pro-His-Cys or Cys-Val-Leu-Cys or Cys-Gly-Tyr-Cys.

In a preferred aspect of the present invention the [Z-B] alignment is repeated at least twice.

In one embodiment the protein variant is extracellularly expressible in large quantities. This is a requirement if the industrial application of the present protein variant is to be feasible. Accordingly, in a preferred embodiment Y is a truncated version of SEQ ID NO:2, thereby increasing the extracellular yield of the protein variant when recombinantly expressed and truncated versions are encompassed by the present invention as functional equivalents of the sequence. Preferably, Y is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ

ID NO:9 or SEQ ID NO:10 or a functional equivalent thereof. In a more preferred embodiment of the invention Y comprises a sequence corresponding to SEQ ID NO:2.

In the context of the invention the term a functional equivalent relates to a sequence processing a corresponding property as the sequences mentioned in the present invention, but wherein one or more amino acids have been substituted with others. Preferably a functional equivalent contains conservative substitutions, i.e. where one or more amino acids are substituted by an amino acid having similar properties, such that a person skilled in the art of protein chemistry will expect the secondary and tertiary structure of the protein to be unchanged. Amino acids suitable for conservative substitutions include those having functionally similar side chains. For example, hydrophobic residues: *e.g.* glycine, alanine, valine, leucine, isoleucine and methionine may replace another such residue. Similarly, conservative substitutions may involve interchanging hydrophilic residues: (*e.g.*: arginine and lysine, glutamine and aspargine, threonine and serine), basic reduces (*e.g.*, lysine, arginine and histidine), and/or acidic residues (*e.g.*, aspartic acid and glutamic acid). Functional equivalents may also, or alternatively, be modified by for example the deletion or addition of amino acids, or the chemical modification of amino acids, as long as the function of the protein is preserved. Furthermore a functional equivalent according to the invention may additionally relate to any truncated sequence having properties identical to the sequences of the invention.

As discussed above it is an object of the present invention to provide a protein variant having an increased reducing activity as compared to a parent PDI, e.g. a wild-type PDI. The reducing activity, i.e. property, may e.g. be more than 25%, 50%, 75%, 100% 150%, 200% or 250% of the parent protein (e.g. wild-type). The specific activity may be increased compared to the parent protein. The specific activity according to the present invention may be defined as activity in insulin reduction units per mg. of protein and may be estimated using the insulin reduction assay described by Bardwell, J. C. A. et al. (Cell 67, pp. 581-589, 1991).

The invention also relates to a fusion polypeptide comprising a polypeptide of the invention and a fusion partner. The fusion protein may e.g. comprise a polypeptide as defined by the invention, and another protein fragment, wherein said other protein fragment is capable of facilitating expression and purification of the polypep-

tide, optionally by reducing the susceptibility of the protein variant to enzymatic degradation. The fusion partner may be a signal peptide, such as, e.g., amino acid number 1-20 of SEQ ID NO:13.

The term polypeptide means a polymer of amino acids and may also be termed protein. The polypeptide may consist of a single polypeptide chain (monomeric) or comprise several associated polypeptides (multimeric, e.g. dimeric). In further embodiments of the invention, the polypeptide of the invention consist of at most 600 amino acids, such as at most 400 amino acids, such as, at most 350, at most 330, at most 300, at most 285, at most 281 amino acids or at most 115 amino acids.

Nucleotide sequences encoding the polypeptides of the invention

The present invention also relates to isolated nucleotide sequences which encode a polypeptide of the present invention. In a preferred embodiment, the nucleotide sequence is SEQ ID NO:14 or SEQ ID NO:16. The present invention also encompasses a nucleotide sequence which encode a polypeptide having the amino acid sequence SEQ ID NO:15, SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15.

The nucleotide sequence of the parent PDI may be isolated or cloned as known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. Thus, the invention relates to a nucleic acid comprising a nucleotide sequence encoding the polypeptide of the invention, the nucleotide sequence may optionally be linked to one or more control sequences that direct the production of the polypeptide in a suitable expression host. The cloning of the nucleotide sequence coding for the parent PDI can be effected, *e.g.*, by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleotide sequence-based amplification (NASBA) may be used.

The nucleotide sequence encoding a parent PDI may be isolated from any cell or microorganism producing the PDI in question, using various methods well known in the art. First, a genomic DNA and/or cDNA library may be constructed using

chromosomal DNA or mRNA from the organism that produces the PDI. Then, if the amino acid sequence of the PDI is known, labelled oligonucleotide probes may be synthesized and used to identify PDI-encoding clones from a genomic library prepared from the organism in question. Alternatively, a labelled oligonucleotide probe containing sequences homologous to another known PDI-gene could be used as a probe to identify PDI-encoding clones, using hybridization and washing conditions of lower stringency. Yet another method for identifying PDI-encoding clones would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming PDI-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a substrate for PDI, thereby allowing clones expressing the PDI to be identified.

Alternatively, the nucleotide sequence encoding the enzyme may be prepared synthetically by established standard methods, *e.g.* the phosphoroamidite method described S.L. Beaucage and M.H. Caruthers, (1981), Tetrahedron Letters 22, p. 1859-1869, or the method described by Matthes et al., (1984), EMBO J. 3, p. 801-805. In the phosphoroamidite method, oligonucleotides are synthesized, *e.g.* in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors. The nucleotide sequence may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof. The nucleotide sequence, e.g. a DNA sequence, may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin or mixed genomic and cDNA origin, prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate, the fragments corresponding to various parts of the entire DNA sequence), in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific primers, for instance as described in US 4,683,202 or R.K. Saiki et al., (1988), Science 239, 1988, pp. 487-491. Alternative methods for providing polypeptides of the invention include gene-shuffling method known in the art including the methods, e.g., described in WO 95/22625 and WO 96/00343.

The introduction of a mutation into the nucleotide sequence to exchange one nucleotide for another nucleotide may be accomplished by site-directed mutagenesis using any of the methods known in the art. Particularly useful is the procedure which utilizes a supercoiled, double stranded DNA vector with an insert of interest and two synthetic primers containing the desired mutation. The oligonucleotide primers, each

complementary to opposite strands of the vector, extend during temperature cycling by means of *Pfu* DNA polymerase. On incorporation of the primers, a mutated plasmid containing staggered nicks is generated. Following temperature cycling, the product is treated with *Dpn*I which is specific for methylated and hemimethylated DNA to digest the parental DNA template and to select for mutation-containing synthesized DNA. Other procedures known in the art may also be used. For a general description of nucleotide substitution, see, *e.g.*, Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95-107. Reference is also made to Morinaga et al., (1984), Biotechnology 2, pp. 646-639. US 4,760,025 disclose the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette. Another method for introducing mutations into PDI-encoding DNA sequences is described in Nelson and Long, (1989), Analytical Biochemistry 180, p. 147-151.

The present invention also relates to nucleic acid constructs comprising a nucleotide sequence of the present invention operably linked to one or more control sequences which direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences. Expression will be understood to include any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

The term nucleic acid construct is synonymous with the term expression cassette when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention.

The term "coding sequence" is defined herein as a nucleotide sequence which directly specifies the amino acid sequence of its protein product. The boundaries of a genomic coding sequence are generally determined by a ribosome binding site (prokaryotes) or by the ATG start codon (eukaryotes) located just upstream of the open reading frame at the 5' end of the mRNA and a transcription terminator sequence located just downstream of the open reading frame at the 3' end of the mRNA. A coding sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleotide sequences.

The term "control sequences" is defined herein to include all components which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide se-

quence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide. The term "operably linked" is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of the nucleotide sequence. The promoter sequence contains transcriptional control sequences which mediate the expression of the polypeptide. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli lac* operon, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis xylA* and *xylB* genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74-94; and in Sambrook *et al.*, 1989, *supra*.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei*

aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleotide sequence encoding the polypeptide. Any terminator that is functional in the host cell of choice may be used in the present invention.

The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleotide sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleotide sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

To allow the secretion of the expressed protein variant, i.e. the extracellular expression, the nucleic acid construct may include a signal sequence, inserted prior to the coding sequence. Thus, the control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a

polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

Effective signal peptide coding regions for filamentous fungal host cells are the signal peptide coding regions obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase.

The constructs further contain one or more exons of the endogenous gene. An exon is defined as a DNA sequence which is copied into RNA and is present in a mature mRNA molecule such that the exon sequence is in-frame with the coding region of the endogenous gene. The exons can, optionally, contain DNA which encodes one or more amino acids and/or partially encodes an amino acid. Alternatively, the exon contains DNA which corresponds to a 5' non-encoding region. Where the exogenous exon or exons encode one or more amino acids and/or a portion of an amino acid, the nucleic acid construct is designed such that, upon transcription and splicing, the reading frame is in-frame with the coding region of the endogenous gene so that the appropriate reading frame of the portion of the mRNA derived from the second exon is unchanged.

The splice-donor site of the constructs directs the splicing of one exon to another exon. Typically, the first exon lies 5' of the second exon, and the splice-donor site overlapping and flanking the first exon on its 3' side recognizes a splice-acceptor site flanking the second exon on the 5' side of the second exon. A splice-acceptor site, like a splice-donor site, is a sequence which directs the splicing of one exon to

another exon. Acting in conjunction with a splice-donor site, the splicing apparatus uses a splice-acceptor site to effect the removal of an intron.

Expression of the polypeptides

The present invention also relates to a vector comprising a nucleotide sequence of the invention, including recombinant expression vectors comprising the nucleotide sequence, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites. Alternatively, the nucleotide sequence of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy

metals, prototrophy to auxotrophs, and the like. Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance. The vector may comprise *Aspergillus* selection markers such as amdS, argB, niaD and sC, a marker giving rise to hygromycin resistance, or the selection may be accomplished by co-transformation, *e.g.* as described in WO 91/17243.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question.

More than one copy of a nucleotide sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleotide sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleotide sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleotide sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Host cells

The present invention also relates to recombinant host cells, comprising a nucleotide sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. A vector comprising a nucleotide sequence of the present invention is introduced into a host cell so that the vector may be maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

The host cell may be a unicellular microorganism, *e.g.*, a prokaryote, or a non-unicellular microorganism, *e.g.*, a eukaryote. The host cell may be chosen from mammal, avian, insect or plant cells, or it may be selected from bacteria or fungi.

Useful unicellular cells are bacterial cells such as gram positive bacteria including, but not limited to, a *Bacillus* cell, *e.g.*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis*; or a *Streptomyces* cell, *e.g.*, *Streptomyces lividans* and *Streptomyces murinus*, or gram negative bacteria such as *E. coli* and *Pseudomonas* sp. In a preferred embodiment, the bacterial host cell is a *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, or *Bacillus subtilis* cell. In another preferred embodiment, the *Bacillus* cell is an alkalophilic *Bacillus*.

The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, *e.g.*, Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), using competent cells (see, *e.g.*, Young and Spizizin, 1961, *Journal of Bacteriology* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporation (see, *e.g.*, Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, *e.g.*, Koehler and Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

In a preferred embodiment, the host cell is a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth *et al.*, *In*, *Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth *et al.*, 1995, *supra*, page 171) and all mitosporic fungi (Hawksworth *et al.*, 1995, *supra*).

In a more preferred embodiment, the fungal host cell is a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980).

In an even more preferred embodiment, the yeast host cell is a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell.

In a most preferred embodiment, the yeast host cell is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* or *Saccharomyces oviformis* cell. In another most preferred embodiment, the yeast host cell is a *Kluyveromyces lactis* cell. In another most preferred embodiment, the yeast host cell is a *Yarrowia lipolytica* cell.

In another more preferred embodiment, the fungal host cell is a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

In an even more preferred embodiment, the filamentous fungal host cell is a cell of a species of, but not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, or *Trichoderma*.

In a most preferred embodiment, the filamentous fungal host cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* or *Aspergillus oryzae* cell. In another most preferred embodiment, the filamentous fungal host cell is a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, or *Fusarium venenatum* cell. In an even most preferred embodiment, the filamentous fungal parent cell is a *Fusarium venenatum* (Nirenberg sp. nov.) cell. In another most preferred embodiment, the filamentous fungal host cell is a *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, *In* Abelson, J.N. and Simon, M.I., editors, *Guide to Yeast Genetics and Molecular Biology*, *Methods in Enzymology*, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, *Journal of Bacteriology* 153: 163; and Hinnen *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

Methods of producing the polypeptides of the invention

In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. Polypeptide may be recovered from the medium by conventional procedures including separating the cells from the medium by centrifugation or filtration. If necessary a purification step may be carried out, for example ion exchange chromatography, affinity chromatography or the like. Thus, the

polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, J.C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

The present invention also relates to a transgenic plant, plant part, or plant cell which has been transformed with a nucleotide sequence encoding a polypeptide of the invention so as to express and produce the polypeptide in recoverable quantities. The polypeptide may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the recombinant polypeptide may be used as such for improving the quality of food or feed, *e.g.*, improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, Poa), forage grass such as festuca, lolium, temperate grass, such as Agrostis, and cereals, *e.g.*, wheat, oats, rye, barley, rice, sorghum, and maize (corn). Examples of dicot plants are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family Brassicaceae), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*.

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers. Also specific plant tissues, such as chloroplast, apoplast, mitochondria, vacuole, peroxisomes, and cytoplasm are considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part. Also included within the scope of the present invention are the progeny of such plants, plant parts and plant cells.

The transgenic plant or plant cell expressing a polypeptide of the present invention may be constructed in accordance with methods known in the art. Briefly, the plant or plant cell is constructed by incorporating one or more expression constructs encoding a polypeptide of the present invention into the plant host genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

Conveniently, the expression construct is a nucleic acid construct which comprises a nucleotide sequence encoding a polypeptide of the present invention operably linked with appropriate regulatory sequences required for expression of the nucleotide sequence in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences is determined, for example, on the basis of when, where, and how the polypeptide is desired to be expressed. For instance, the expression of the gene encoding a polypeptide of the present invention may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague *et al.*, 1988, *Plant Physiology* 86: 506.

For constitutive expression, the 35S-CaMV promoter may be used (Franck *et al.*, 1980, *Cell* 21: 285-294). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu *et al.*, 1998, *Plant and Cell Physiology* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad *et al.*, 1998, *Journal of Plant Physiology* 152: 708-711), a promoter from a seed oil body protein (Chen *et al.*, 1998, Plant and Cell Physiology 39: 935-941), the storage protein napA promoter from *Brassica napus*, or any other seed specific promoter known in the art, *e.g.*, as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcs* promoter from rice or tomato (Kyozuka *et al.*, 1993, *Plant Physiology* 102: 991-1000, the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Molecular Biology* 26: 85-93), or the *aldP* gene promoter from rice (Kagaya *et al.*, 1995, *Molecular and General Genetics* 248: 668-

674), or a wound inducible promoter such as the potato pin2 promoter (Xu *et al.*, 1993, *Plant Molecular Biology* 22: 573-588).

A promoter enhancer element may also be used to achieve higher expression of the enzyme in the plant. For instance, the promoter enhancer element may be an intron which is placed between the promoter and the nucleotide sequence encoding a polypeptide of the present invention. For instance, Xu *et al.*, 1993, *supra* disclose the use of the first intron of the rice actin 1 gene to enhance expression.

The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including *Agrobacterium*-mediated transformation, virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

Presently, *Agrobacterium tumefaciens*-mediated gene transfer is the method of choice for generating transgenic dicots (for a review, see Hooykas and Schilperoort, 1992, *Plant Molecular Biology* 19: 15-38). However, it can also be used for transforming monocots, although other transformation methods are generally preferred for these plants. Presently, the method of choice for generating transgenic monocots is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667-674). An alternative method for transformation of monocots is based on protoplast transformation as described by Omirulleh *et al.*, 1993, *Plant Molecular Biology* 21: 415-428.

Following transformation, the transformants having incorporated therein the expression construct are selected and regenerated into whole plants according to methods well-known in the art.

The present invention also relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a transgenic plant or a plant cell comprising a nucleotide sequence encoding a polypeptide capable of reducing disulfide bonds of the present invention under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

Properties of the polypeptides of the invention

The polypeptides of the invention preferably have an improved capability of breaking protein disulfide bonds. Of particular interest is a polypeptide, capable of being produced in large quantities and having a redox potential which is lower than that of SEQ ID NO:13 (e.g. without the signal peptide) or which is lower than that of a subsequence of SEQ ID NO:13. In other embodiment, the polypeptide of the invention has a redox potential which is lower than that of the amino acid sequence of amino acid number 1-281 of SEQ ID NO:13 or which is lower than that of the amino acid sequence of amino acid number 21-281 of SEQ ID NO:13.

In one embodiment the polypeptide of the invention has a redox potential which is decreased with at least 10mV, at least 20mV, at least 30mV or at least 50mV compared to the protein disulfide isomerase having the amino acid sequence SEQ ID NO:13 (e.g. without the signal peptide) or a subsequence of SEQ ID NO:13, such as amino acid number 21-281 of SEQ ID NO:13. In further embodiments the polypeptide of the invention has a redox potential of less than –200mV, such as less than –220mV, less than –240mV, or less than –250mV. In one embodiment wherein the polypeptide is a variant of a parent PDI, the variant polypeptide has a redox potential which is lower than that of the parent disulfide isomerase, such as e.g. at least, 10%, or at least 20% lower. The lower redox potential means that the polypeptide has higher reducing properties than that of the parent disulfide isomerase. The polypeptide of the invention may have a redox potential which is e.g. at least 30 mV or at least 50mV lower as compared to the redox potential of the parent protein disulfide isomerase, such as, e.g. a redox potential which is in the range of 30-80 mV or 30-50mV lower as compared to the redox potential of the parent protein disulfide isomerase .

The redox potential may be determined as described in J. Lundstrom, et al (1992) J. Biol. Chem. 267, 9047 – 9052.

The redox potential of an enzymes capable of reducing disulfide bonds is a measure of whether the reaction has equilibrium to the right (higher redox potential) or to the left (lower redox potential) in the following reaction:

$$Cys\text{-}Xxx\text{-}Yyy\text{-}Cys = Cys\text{-}Xxx\text{-}Yyy\text{-}Cys + 2e^- + 2H^+ .$$

\ / \ /
SH HS S - S

The following table disclose standard redox potentials for the reaction.

Molecule	Source	Sequence	Redox potential	Reference
PDI	Bovine	CGHC	-190 mV	(1)
Thioredoxin (Trx)	*E. coli*	CGPC	-270 mV	(2)
P34H Trx	*E. coli*	CGHC	-235 mV	(2)
Thioredoxin reductase	*E. coli*	CATC	-250 mV	(3) (4)
DsbA	*E. coli*	CPHC	-89 mV	(5)
CaBP1	Rat	CGHC	(about -200 mV)	(2)

(1) Lundström J. and A. Holmgren (1993) Biochemistry 32, 6649 -6655.

(2) Lundström-Ljung, J. et. al (1995) FEBS Letters 357, 305 - 308.

(3) Russel M and P Model (1988): J. Biol. Chem. 263, 9015-9019.

(4) Siedler, F. et al (1993) Biochemistry 32, 7488-7495.

(5) Wunderlich M. and R. Glockshuber (1993): Protein Science 2, 717 -726

Use of the polypeptide of the invention

The present invention may be applied to a number of industrial fields. Allergy towards certain food items is an increasing concern to many people. Accordingly, the polypeptides of the present invention may be applied for the use as an allergen reducing agent.

In particular, the polypeptides may be used as an allergen reducing agent in food, such as gluten or milk. By applying the polypeptides of the present invention to various food or feed products before, during, or after fabrication, disulfide bonds will be reduced and thus an allergic reaction in susceptible individuals avoided.

Another area of particular interest in relation to the present invention is the increasing allergy in people, i.e. infants, toward milk supplements. A common practice of destabilising allergy promoting proteins in milk is that of heat treatment. However, heat treatment may reduce only some of the allergens, and has the unfortunate side effect of at the same time reducing the nutritional value of the milk. By using the polypeptides disclosed by the present invention it is possible to increase the suscep-

tibility of milk proteins, such as α-lactalbumin, to the degradation by the enzyme trypsin without having to heat treat the milk under high temperature conditions. The polypeptides may be used for reducing allergens in food or feed, e.g. gluten or milk based products, including beverages, such as infant formula and dietary drinks.

It therefore follows that the present invention presents an advantage of reducing allergens in milk and at the same time preserve its nutritional value, in addition to having the benefit of being manufactured on a large industrial scale.

In another aspect of the present invention a preferred embodiment is applying the polypeptides of the invention to the baking industry for the reduction of allergens in gluten. Like milk allergies many people suffer from gluten intolerance and this fact signifies yet another vast industrial application of the invention.

Further aspects is use of the polypeptides of the invention for increasing the digestibility of food or feed, such as, e.g. for increasing the digestibility of milk or wheat based food or feed products. The polypeptides may also be used in the manufacturing of a cosmetic product and contemplated are also cosmetic products comprising a polypeptide of the invention.

Yet another aspect of the present invention is the use of the polypeptides for the treatment or degradation of scleroproteins, the treatment and cleaning of fabrics, additives to detergents and pharmaceutical preparations for the treatment of eye sufferings.

According to the invention it is further envisaged that the gene of the polypeptides of the present invention may be expressed in plants or animals for the pretreatment of allergens before processing the plants or animals into products.

The present invention also relates to compositions comprising the polypeptide according to the invention. The content of the polypeptide per gram of composition depends on the use of the composition. However, the compositions may suitably comprise 0,01 – 1,00 mg of polypeptide per g, preferably 0,05 – 0,1 mg of polypeptide per gram.

Further to the invention the composition comprising the polypeptide of the invention may have any suitable form, such as the form of a granulate, a stabilised liquid or a protected enzyme. The composition may also comprise a suitable redox partner, such as an organic or inorganic reductant.

For many of the applications mentioned above it may be convenient that the composition may comprise at least one other enzyme than the polypeptide of the invention, such as a protease, an amylase, a lipase, a hydrolase, a peroxidase, a cellulase, a transglutaminase, a glucose oxidase, a xylanase, a pectin methyl esterase or a pectin lyase. In particular, in connection with reducing allergens in milk the composition may comprise other enzymes, such as proteases and hydrolases.

The invention also relates to food additives or cosmetic products comprising a polypeptide of the invention.

The present invention is further described by the following examples which should not be construed as limiting the scope of the invention.

EXAMPLES

The following are examples of the construction of a reducing protein variant truncation.

EXAMPLE 1: Construction of the reducing variant H61P PDI truncation

a) Construction of the expression plasmid pCaHj 527:

The *Aspergillus oryzae* expression plasmid pCaHj 483 (cf. WO 98/00529) consists of an expression cassette based on the Aspergillus niger neutral amylase II promoter fused to the *Aspergillus nidulans* triose phosphate isomerase non translated leader sequence (Pna2/tpi) and the *Aspergillus niger* amyloglycosidase terminater (Tamg). Also present on the plasmid is the *Aspergillus* selective marker *amdS* from *Aspergillus nidulans* enabling growth on acetamide as sole nitrogen source. These elements are cloned into the *E. coli* vector pUC19. The ampecillin resistance marker enabling selection in *E. coli* of this plasmid was replaced with the URA3 marker of *Saccharomyces cerevisiae* that can complement a *pyrF* mutation in *E. coli* in the following way:

The pUC 19 origin of replication was PCR amplified from pCaHj483 with the primers:

142779: TTG AAT TGA AAA TAG ATT GAT TTA AAA CTT C (SEQ ID NO:19)

142780: TTG CAT GCG TAA TCA TGG TCA TAG C (SEQ ID NO:20)

The primer 142780 introduces a Bbu I site in the PCR fragment.

The Expand PCR system (Roche Molecular Biochemicals, Basel, Schweitzerland) was used for the amplification following the manufacturers instructions for this and the subsequent PCR amplifications.

The URA3 gene was amplified from the general *S. cerevisiae* cloning vector pYES2 (Invitrogen corporation, Carlsbad, Ca, USA) using the primers:

140288: TTG AAT TCA TGG GTA ATA ACT GAT AT (SEQ ID NO:21)

142778: AAA TCA ATC TAT TTT CAA TTC AAT TCA TCA TT (SEQ ID NO:22)

The primer 140288 introduces an EcoR I site in the PCR fragment.

The two PCR fragments were fused by mixing them and amplifying using the primers 142780 and 140288 using the splicing by overlap method (Horton et. al, (1989), Gene, 77, 61-68). The resulting segment was digested by EcoR I and Bbu I and ligated to the largest fragment of pCaHj 483 being digested by the same enzymes. The ligation mixture was used to transform the *pyrF⁻ E. coli* strain DB6507 (ATCC 35673) made competent by the method of Mandel and Higa (Mandel, M. and Higa, A. (1970), J. Mol. Biol. 45, 154). Transformants were selected on solid M9 medium (Sambrook et. al (1989) Molecular cloning, a laboratory manual, 2. edition, Cold Spring Harbor Laboratory Press) supplemented with 1 g/l casamino acids, 500 µg/l thiamine and 10 mg/l kanamycin.

A plasmid from such a transformant was called pCaHj 527 and is outlined in Figure 1.

b) Construction of a truncated PDI:

This section "b)" is also disclosed in WO 95/00636 (Novo Nordisk A/S).

Cloning of *Aspergillus oryzae* and *Aspergillus niger* PDI encoding genes

1.1. Design of oligo nucleotides for PCR amplification

PDI from different organisms are highly homologous especially near the active site residues. In Fig.7, the following 7 PDI gene products were aligned:

Bovine (Bos taurus) PDI (Yamauchi et al., Biochem. Biophys. Res. Commun. 146:1485-1492, 1987), Chicken (Gallus gallus) PDI (Parkkonen et al., Biochem. J. 256:1005-1011, 1988), Human (Homo sapiens) PDI (Rapilajaniemi et al. EMBO J. 6:643-649,

1987), Mouse (Mus musculus) PDI (Gong, et al., Nucleic Acids Res. 16:1203, 1988), Rabbit (Oryctolagus cuniculus) PDI (Fliegel et al., J. Biol. Chem. 265:15496-15502, 1990), Rat (Rattus norvegicus) PDI (Edman et al., Nature 317:267-270, 1985), Yeast (Saccharomyces cerevisiae) PDI (Tachikawa et al., J. Biochem. 110:306-313).

Each subunit contains two active centres (Freedman et al., Cell 57:1069-1072, 1989) and the homology in the surroundings of these active centres are particularly strong. A consensus amino acid sequence for the active centre closest to the N-terminus was determined from the alignment as -APWCGHCK-, and an oligo deoxyribonucleotide encoding the peptide -WCGHCK- and extended with an EcoRI site in the 5` end, was synthesized:

5`TGGAATTCTGGTGYGGNCAYTGYAA3` (primer 4762, 25 nucleotides, 32 species) (Y=C or T; R=A or G; N=A, T, C, or G).

A consensus amino acid sequence for the active centre closest to the C-terminus was determined: -YAPWCGHCK-, and an oligo deoxyribonucleotide encoding the peptide -YAPWCG- in antisense and extended with a BamHI site in the 5` end was synthesized:

5`TGGGATCCRCACCANGGNGCRTA3` (primer 4763, 23 nucleotides, 64 species).

These oligo deoxyribonucleotides (primers 4762 and 4763) were used as primers in a PCR reaction to amplify PDI-encoding gene fragments from genomic DNA from *A. oryzae* and *A. niger*.

1.2 Amplification and cloning of fragments of PDI-encoding genes.

Genomic DNA was prepared from *Aspergillus oryzae* IFO 4177 and *Aspergillus niger* A524 as described by Yelton et al. (Proc. Natl. Acad. Sci. USA 81:1470-1474, 1984).

PCR reaction mixtures contained Taq DNA polymerase buffer supplied by Clontech laboratories Inc. and diluted as described, 250 µM of each of dATP, dCTP, dGTP, and, dTTP, 100 pmol of each of primers 4762 and 4763, and 0.5 µg of genomic DNA of either *A. niger* or *A. oryzae*. The total reaction volume was 0.1 ml, and it was covered with 0.05 ml paraffin oil.

The following program was run on a Cetus Perkin Elmer thermal cycler:

1. cycle: 94°C for 2 min., (when the temperature reached 94°C 2.5 U of Taq DNA polymerase supplied by Clontech laboratories Inc. was added).
10 cycles: 94°C for 1 min., 50°C for 1 min., and 72°C for 2 min.
30 cycles: 94°C for 1 min., 55°C for 1 min., and 72°C for 2 min.
1 cycle: 72°C for 5 min.

The reaction mixtures were loaded on an agarose gel, and both the *A. oryzae* and the *A. niger* DNA produced fragments of approximately 1.1 kb.

The fragments were digested with EcoRI and BamHI and ligated to pUC19 (Yanisch-Perron et al., Gene 33:103-119, 1985). The ligation mixture was transformed into *E. coli* DH5αF` (Woodcock et al., Nucleic Acids Res. 17:3469-3478). Recombinant plasmids were subjected to sequence analysis using the Sequenase™ kit (United States Biochemical) and a M13 universal primer following the manufacturers instructions. The analysis confirmed that both in the case of *A. oryzae* and in that of *A. niger* sequences homologous to other PDI genes were amplified and cloned.

1.3 Genome cloning of the *A. oryzae* PDI-encoding gene.

Genomic DNA from *A. oryzae* was digested with the following restriction enzymes supplied by New England Biolabs Inc.: HindIII, BamHI, BamHI+HindIII, EcoRI, EcoRI+HindIII, SalI, SalI+HindIII, BglII, BglII+HindIII, PstI and PstI+HindIII. After digestion, the reaction mixtures were run on a 1% agarose gel and then blotted onto an Immobilon N™ membrane (Millipore Corporation) following the manufacturers instructions. The membrane was probed with the cloned *A. oryzae* PCR product isolated as a BamHI-EcoRI fragment and radio labelled with ^{32}P. After stringent washes the membrane was subjected to autoradiography.
Genomic DNA from *A. niger* was digested with the following restriction enzymes: BglII, BamHI, BamHI+BglII, EcoRI, EcoRI+BglII, SalI, SalI+BglII, HindIII, HindIII+BglII, PstI and PstI+BglII. The Southern blot was made as described with *A. oryzae*, only the *A. niger* PCR product was used as probe.

1.4 Construction of genomic *A. oryzae* library.

Southern analysis indicated that the *A. oryzae* PDI gene was located on a 2.3 kb BamH I, Hind III fragment. Genomic *A. oryzae* DNA was digested with BamH I and Hind

III and fragments ranging from 1.9 - 3 kb were isolated from an agarose gel. This mixture of fragments was ligated to pUC19 digested with BamHI and Hind III. The ligation mixture was used to transform *E. coli* DH5αF`. The transformed *E. coli* cells were spread onto 10 agar plates using ampicillin selection.

1.5 Screening of the *A. oryzae* genomic library.

The libraries were screened using the filter colony hybridization method described by Gergen et al. (Nucleic Acids Res. 7:2115-2136, 1979). The probe that was used for the Southern blot was also used for the colony hybridization. Positive clones were isolated and confirmed by sequence analysis using sequencing primers designed from the sequences of the PDI fragments. One of the plasmids containing the desired fragment was termed pCaHj 425.

1.6 Sequence of the gene.

The gene was sequenced using the Tag DyeDeoxy™ Terminator cycle sequencing kit supplied by Applied Biosystems following the manufacturer`s instructions. The sequence reactions were run on an Applied Biosystems 373A DNA sequencer and the data were evaluated using the Macintosh computer program SegEd version 1.0 supplied by Applied Biosystems.

The sequence of the *A. oryzae* PDI gene is shown in SEQ ID NO: 12 and the corresponding amino acid sequence is shown in SEQ ID NO:13.

Expression of a truncated form of the *A. oryzae* PDI gene.

Construction of expression plasmids.

The PDI gene of *A. oryzae* was truncated by introduction of a stop codon. This was done by PCR amplification of the PDI gene using a 5' PCR primer harbouring a BamH I site at its 5' end and a 3' primer corresponding to a truncation harbouring a Hind III site.

The sequence of the 5' primer was:

5' TTCGGATCCACCATGCGGACTTTCGCACC 3' 5205.

The sequence of the 3' primers was:

5' CCAAGCTTAGTGTTTCTCGGCGATGAACTT 3' 6314.

Primer 6314 introduced a stop codon after aminoacid 281.

The expression plasmid were constructed by PCR amplification using primer 5205 in combination with 6314 and pCaHj 425 as template using standard PCR conditions. The generated PCR segment was digested with BamH I and Hind III and inserted into pMTH 1560 (cf. WO 98/00529) digested with the same enzymes. The constructed plasmid was named pCaHj 445 (from primer 6314).

c) Construction of the H61P mutated truncated PDI (expression plasmid pCaHj 548):

The truncated PDI of pCaHj 445 was inserted into pCaHj 527 by the digestion of pCaHj 445 by BamH I and Xho1, and ligating the PDI fragment to pCaHj 527 digested by the same enzymes. The ligation mixture was transformed into *E. coli* DB6507 as described above to form the plasmid pCaHj 545.

pCaHj 545 was then used as template for a PCR reaction with the following primers:

160712: CCC TTG CAA GGC TCT CGC TCC (SEQ ID NO:23)

18699: TTG CCC TCA TCC CCA TCC TTT (SEQ ID NO:24)

160712 was phosphorylated in the 5' end. The primer covers the active site of PDI and alters Histidine in position 61 to proline.

The formed PCR segment was digested by Xho I and ligated to the large fragment of pCaHj 545 being digested by Bal I and Xho I. The ligation mixture was used to transform *E. coli* DB6507 as described above to form the plasmid pCaHj 548 (coding for amino acid sequence SEQ ID NO:15 (the H61P mutated truncated PDI), thus leading to the polypeptide with an amino acid sequence 21-281 of SEQ ID NO:15. The construction of the plasmid is outlined in Figure 2.

d) Transformation into *Aspergillus oryzae*:

pCaHj 545 (comprising a nucleotide sequence coding for the amino acid sequence of amino acid number 1 to 281 of SEQ ID NO:13) and pCaHj 548 (comprising SEQ ID NO:14 coding for the amino acid sequence SEQ ID NO:15) were transformed into the protease weak *Aspergillus oryzae* strain JaL228 (WO 98/12300), fermented and recovered as described in WO 95/00636.

EXAMPLE 2: Treatment of α-lactalbumin with a protein variant of PDI

The following experiments were performed to establish a simple and alternative method for reducing the allergenicity of food allergens by the reduction of disulfide bonds. This was done by assessing the capacity of protein disulfide isomerases to destabilise α-lactalbumin and to increase the susceptibility of this protein to proteolysis *in vitro*. As a model substrate the cow's milk allergen α-lactalbumin having 4 disulfide bonds was chosen.

Materials and methods

Substrate:

250 µl 1M K_3PO_4, pH 6.5

250 µl [5, 10 or 20 mg/ml] α-lactalbumin

10 µl 100 mM β-mercapto-ethanol

Enzyme solutions:

PDI concentrations ranging from 0.04 - 4.00 mg/ml for:

a) recombinant fungal protein disulfide isomerase #960112/BRNi, and

b) DsbA 9412.206 22.06-94 BRNi/Bgra

Trypsin stock solution [5 mg/ml]

Reaction conditions:

25 µl H_2O or enzyme solution was added to 25 µl substrate. Trypsin was added in a trypsin: α-lactalbumin ratio of 1:25. This reaction mixture was incubated overnight at 37 °C.

Analysis conditions:

20 µl Laemmli sample buffer, including 120 mM DTT was added to 20 µl reaction mixture. After 10 min at 70 °C, 10 µl 1M iodoacetamide was added. Electroporesis was performed at 125 V for 110 min on a 4-20% acrylamide gradient gel (NOVEX Pre-Cast Gels). Electroblotting to nitrocellulose membranes was performed at 30 V for 90 min. Protein staining was done as described in F-9503186.

Results

The results (Figure 5 and Figure 6) show that treatment of α-lactalbumin with a recombinant PDI, but not DsbA, destabilises α-lactalbumin and makes it more susceptible to trypsin action under the current conditions for proteolysis. When 5 mg/ml

α-lactalbumin was incubated with 0.20 mg/ml PDI, all α-lactalbumin was digested by trypsin.

Optimisation of the 'PDI:substrate' ratio reveals that all α-lactalbumin can be digested by trypsin when there is no more than a 50 X excess (in mg) as compared to PDI. Accordingly the ratio is 1 mole of PDI (MW = 90 kDa) per 300 moles of α-lactalbumin (MW = 15 kDa).

From the above it can be concluded that recombinant fungal protein disulfide isomerase has a significant effect on the stability of the cow's milk allergen α-lactalbumin.

CLAIMS

1. A polypeptide capable of reducing disulfide bonds, the polypeptide comprising an amino acid sequence selected from the group consisting of:

(i) an amino acid sequence having at least 60% identity with the amino acid sequence set forth in SEQ ID NO:15 or amino acid number 21-281 of SEQ ID NO:15;

(ii) an amino acid sequence having at least 60% similarity with the amino acid sequence set forth in SEQ ID NO:15 or amino acid number 21-281 of SEQ ID NO:15;

(iii) an amino acid sequence having at least 60% identity with the amino acid sequence set forth in SEQ ID NO:17; and

(iii) an amino acid sequence having at least 60% similarity with the amino acid sequence set forth in SEQ ID NO:17;

provided that a position in the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is selected from the group consisting of:

i) Cys-Gly-Pro-Cys,

ii) Cys-Ala-Thr-Cys,

iii) Cys-Val-Leu-Cys, and

iv) Cys-Gly-Tyr-Cys.

2. The polypeptide of claim 1, wherein the comprised amino acid sequence has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence set forth in SEQ ID NO:15, amino acid number 21-281 of SEQ ID NO:15 or SEQ ID NO:17.

3. The polypeptide according to claim 1, which polypeptide has Cys-Gly-Pro-Cys in the position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

4. The polypeptide of claim 3, wherein the comprised amino acid sequence has at at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or

similarity with the amino acid sequence set forth in SEQ ID NO:15, SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15.

5. A polypeptide capable of reducing disulfide bonds, which polypeptide is encoded by a nucleotide sequence which hybridizes under low, medium or high stringency conditions with (i) SEQ ID NO:14; (ii) SEQ ID NO:16; (iii) a subsequence of (i) or (ii) of at least 100 nucleotides; or (iv) a complementary strand of (i), (ii) or (iii); provided that a position of the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is selected from the group consisting of:

i) Cys-Gly-Pro-Cys,
ii) Cys-Ala-Thr-Cys,
iii) Cys-Val-Leu-Cys, and
iv) Cys-Gly-Tyr-Cys.

6. The polypeptide according to claim 5, which polypeptide has Cys-Gly-Pro-Cys in the position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

7. The polypeptide according to any of the preceding claims, wherein the comprised amino acid sequence is SEQ ID NO:15 or SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15.

8. The polypeptide according to any of the preceding claims, which polypeptide consists of the amino acid sequence SEQ ID NO:15 or SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15.

9. A polypeptide capable of reducing disulfide bonds, which polypeptide

(i) is encoded by the nucleotide sequence contained in plasmid pCaHj548.

(ii) comprises an amino acid sequence having at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence encoded by the nucleotide sequence contained in plasmid pCaHj548; or

(iii) consist of an amino acid sequence having at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity or similarity with the amino acid sequence encoded by the nucleotide sequence contained in plasmid pCaHj548,

provided that a position of the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO: 13 is selected from the group consisting of:

i) Cys-Gly-Pro-Cys,

ii) Cys-Ala-Thr-Cys,

iii) Cys-Val-Leu-Cys, and

iv) Cys-Gly-Tyr-Cys.

10. A polypeptide capable of reducing disulfide bonds, the polypeptide consisting of an amino acid sequence having:

(i) at least 60% identity or similarity with the amino acid sequence set forth in SEQ ID NO:15 or with amino acid number 21-281 of SEQ ID NO:15; or

(ii) at least 60% identity or similarity with the amino acid sequence set forth in SEQ ID NO:17;

provided that a position in the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is selected from the group consisting of:

i) Cys-Gly-Pro-Cys,

ii) Cys-Ala-Thr-Cys,

iii) Cys-Val-Leu-Cys, and

iv) Cys-Gly-Tyr-Cys.

11. The polypeptide according to claim 10, which polypeptide has Cys-Gly-Pro-Cys in the position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

12. The polypeptide of claim 10 or 11, wherein the amino acid sequence has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identity or simi-

larity with the amino acid sequence set forth in SEQ ID NO:15 or SEQ ID NO: 17 or amino acid number 21-281 of SEQ ID NO:15.

13. A polypeptide capable of reducing disulfide bonds, which polypeptide is a variant of a polypeptide having the amino acid sequence of SEQ ID NO:15, SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15; the variant comprising substitution(s), deletion(s), and/or insertion(s) of one or more amino acids; provided that the variant comprises an amino acid sequence selected from the group consisting of:

i) Cys-Gly-Pro-Cys,

ii) Cys-Ala-Thr-Cys,

iii) Cys-Val-Leu-Cys, and

iv) Cys-Gly-Tyr-Cys,

in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

14. The polypeptide according to claim 13, which polypeptide has Cys-Gly-Pro-Cys in the position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

15. The polypeptide of claim 13 or 14, having an amino acid sequence having

(i) at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity or similarity with SEQ ID NO:15 or SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15; or

(ii) 100% similarity with SEQ ID NO:15 or SEQ ID NO:17 or amino acid number 21-281 of SEQ ID NO:15.

16. The polypeptide according to any of the preceding claims further comprising an amino acid sequence which has at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identity or similarity with the amino acid sequence set forth in SEQ ID NO:18 provided the position corresponding to amino acid residues numbered 31-34 in SEQ ID NO:18 is selected from the group consisting of:

(i) Cys-Gly-Pro-Cys,

(ii) Cys-Ala-Thr-Cys,

(iii) Cys-Val-Leu-Cys,

(iv) Cys-Gly-Tyr-Cys, and

(v) Cys-Gly-His-Cys

17. A polypeptide capable of reducing disulfide bonds, the polypeptide being a variant of a parent protein disulfide isomerase by having an amino acid sequence differing from that of the parent protein disulfide isomerase in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 , which position is Cys-Gly-Pro-Cys for the variant polypeptide.

18. The polypeptide according to claim 17 having substitution(s), deletion(s), and/or insertion(s) of one or more amino acids compared to the parent protein disulfide isomerase, provided that the polypeptide comprises the amino acid sequence Cys-Gly-Pro-Cys in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

19. The polypeptide of claim 18, wherein the amino acid sequence of the polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity or similarity with the amino acid sequence of the parent protein disulfide isomerase.

20. The polypeptide of any of claims 17-19, wherein the parent protein disulfide isomerase is a PDI from *A. oryzae* or a subsequence thereof, SEQ ID NO: 13, or a subsequence of SEQ ID NO:13 having protein disulfide isomerase activity, such as, e.g. amino acid no. 21-115 of SEQ ID NO: 13, e.g. amino acid no. 1-281 or 21-281 of SEQ ID NO: 13.

21. The polypeptide of any of claims 17-20, wherein the parent protein disulfide isomerase is encoded by a nucleotide sequence which hybridises under low, medium or high stringency conditions with (i) SEQ ID NO:14 or SEQ ID NO:16; (ii) nucleotides 131 to 365 of SEQ ID NO:12); (iii) a subsequence of (i) or (ii) of at least 100 nucleo-

tides; or (iv) a complementary strand of (i), (ii) or (iii); provided that the position of the polypeptide corresponding to amino acid residues numbered 58-61 in SEQ ID NO: 13 is Cys-Gly-Pro-Cys.

22. A method for providing a polypeptide capable of reducing disulfide bond, the method comprising a step selected from the group consisting of:

(a1) altering at least one amino acid residue in a parent protein disulfide isomerase by amending the amino acid X_1 to Glu and/or the amino acid Y_2 to Pro in the site corresponding to Cys-X_1-Y_2-Cys of the parent protein disulfide isomerase (PDI),
to obtain a variant comprising Cys-Gly-Pro-Cys;

(a2) altering at least one amino acid residue in a parent protein disulfide isomerase by amending the amino acid X_1 to Ala and/or the amino acid Y_2 to Thr in a site corresponding to Cys-X_1-Y_2-Cys of the parent protein disulfide isomerase (PDI),
to obtain a variant comprising Cys-Ala-Thr-Cys;

(a3) altering at least one amino acid residue in a parent protein disulfide isomerase by amending the amino acid X_1 to Val and/or the amino acid Y_2 to Leu in the site corresponding to Cys-X_1-Y_2-Cys of the parent protein disulfide isomerase (PDI),
to obtain a variant comprising Cys-Val-Leu-Cys; and

(a4) altering at least one amino acid residue in a parent protein disulfide isomerase by amending the amino acid X_1 to Gly and/or the amino acid Y_2 to Tyr in the site corresponding to Cys-X_1-Y_2-Cys of the parent protein disulfide isomerase (PDI),
to obtain a variant comprising Cys-Gly-Tyr-Cys.

23. The method according to claim 22, comprising the step (a1) and wherein Cys-Gly-Pro-Cys is in a position corresponding to amino acid number 58-61 of SEQ ID NO:13.

24. The method of claim 23, further comprising the step (b) of substitution(s), deletion(s), and/or insertion(s) of one or more amino acids compared to the parent protein

disulfide isomerase, provided that the produced polypeptide comprises the amino acid sequence Cys-Gly-Pro-Cys in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13.

25. The method of any of claims 22-24, wherein the polypeptide has an amino acid sequence which has at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% identity or similarity with the amino acid sequence of the parent protein disulfide isomerase.

26. A polypeptide capable of reducing disulfide bonds obtainable by the method of any of claims 22-25.

27. A fusion polypeptide comprising a polypeptide according to any of the preceding claims and a fusion partner.

28. The fusion polypeptide according to claim 27, wherein said other protein fragment is capable of facilitating expression and purification of the polypeptide.

29. The fusion polypeptide according to claim 28, wherein the fusion partner is amino acid number 1-20 of SEQ ID NO:15.

30. The polypeptide according to any of the preceding claims capable of being expressed extracellularly in a microbial host organism.

31. The polypeptide according to any of claims 1-21 or any of claims 26-30 having a redox potential which is decreased with at least 10mV, at least 20mV, at least 30mV compared to the protein disulfide isomerase having the amino acid sequence SEQ ID NO:13 or amino acid number 21-281 of SEQ ID NO:13.

32. The polypeptide according to any claims 1-21 or any of claims 26-30 having a redoxpotential of less than –200mV.

33. The polypeptide according to any of the claims 17-20 or 26 having a redox potential which is lower than that of the parent disulfide isomerase.

34. The polypeptide according to claim 33 having a redox potential which is at least 30 mV lower as compared to the redoxpotential of the parent protein disulfide isomerase.

35. The polypeptide according to any of claims 1-21 or 26-30 having a redoxpotential which is lower than that of (i) SEQ ID NO:13 without amino acid number 1-20 or (ii) the amino acid sequence given by amino acid number 21-281 of SEQ ID NO:13.

36. The polypeptide according to any of the preceding claims, which polypeptide is a protein disulfide isomerase (PDI, EC 5.3.4.1.).

37. A polypeptide which is a protein disulfide isomerase variant having increased reducing properties when compared to the wild-type protein, with the formula of:

$X - A - [Z - B]_n - Y$, wherein

A is the amino acid sequence: Cys-Gly-Pro-Cys,

X is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:1 or a functional equivalent thereof,

Y is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:2 or a functional equivalent thereof,

Z is an amino acid sequence comprising a sequence corresponding to SEQ ID NO:11 or a functional equivalent thereof.

B is individually selected from the amino acid sequence Cys-Gly-Pro-Cys or Cys-Gly-His-Gly or Cys-Ala-Thr-Cys or Cys-Pro-His-Cys or Cys-Val-Leu-Cys or Cys-Gly-Tyr-Cys,

n = 0 or an integer from 1 to 3

38. The protein variant according to claim 37, wherein Y is an amino acid sequence selected from the group consisting of SEQ ID NO:3; SEQ ID NO:4; SEQ ID NO:5; SEQ ID NO:6; SEQ ID NO:7; SEQ ID NO:8; SEQ ID NO:9; and SEQ ID NO:10.

39. The polypeptide according to claim 37 or 38, wherein n is 1.

40. The polypeptide according to claim 37 or 38, wherein n is 0.

41. The polypeptide according to claim 40, which polypeptide has an amino acid sequence having at least 60%, at least 65%, least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity or similarity with the amino acid sequence shown as amino acids 21-281 of SEQ ID NO:13.

42. A polypeptide which is a protein disulfide isomerase having an amino acid sequence which in a position corresponding to amino acid residues numbered 58-61 in SEQ ID NO:13 is Cys-Gly-Pro-Cys.

43. A nucleic acid comprising a nucleotide sequence encoding the polypeptide of any of the preceding claims.

44. The nucleic acid according to claim 43 which is DNA.

45. A nucleotide sequence encoding a polypeptide as defined in any of the preceding claims.

46. A nucleic acid construct comprising the nucleotide sequence of claim 45 operably linked to one or more control sequences that direct the production of the polypeptide in a suitable expression host.

47. A vector comprising the nucleotide sequence of claim 45.

48. The vector according to claim 47 which is a recombinant expression vector.

49. The vector according to claim 47 or 48 which is a plasmid.

50. A transformed cell harbouring the vector of any of claims 47-49.

51. The transformed cell according to claim 50 capable of expressing the polypeptide.

52. A recombinant host cell comprising the nucleic acid construct of claim 46, which cell is capable of expressing the encoded polypeptide.

53. The cell according to any of claims 50-52, which is a microbial cell

54. The cell according to claim 53 which is a bacterial or a fungal cell.

55. The cell according to claim 54, which cell is a gram-positive bacterium, e.g. of the genus *Bacillus* or *Streptomyces* or a cell of a gram-negative bacterium, e.g. of the genus *Escherichia*.

56. The cell according to claim 54, which cell is a yeast cell, e.g. of the genus *Saccharomyces*, or a cell of a filamentous fungus, e.g. of the genus *Aspergillus* or *Fusarium*.

57. The cell according to claim 56, which cell is an *Aspergillus* species selected from the group consisting of *A. oryzae*, *A. niger* and *A. nidulans*.

58. A method of producing a polypeptide according to any of claims 1-21 or 26-42, wherein a host cell containing an expression cassette comprising a nucleotide sequence encoding the polypeptide is cultured in a suitable medium under conditions promoting the expression of the protein.

59. A method of producing a polypeptide according to any of claims 1-21 or 26-42, comprising inserting a nucleic acid according to claim 40 into a vector which is able to replicate in a host cell, introducing the resulting recombinant vector into the host cell, culturing the host cell in a culture medium under conditions sufficient to effect expression of the polypeptide, and recovering the polypeptide from the host cell or the culture medium; or synthesising the polypeptide by solid or liquid phase synthesis.

60. A method of producing the polypeptide of any of claims 1-21 or 26-42, which method comprises the step of (i) cultivating the cell of any of claims 47-54 so as to express and optionally secrete the polypeptide; and (ii) recovering the polypeptide.

61. The method according to any of claims 58-60, in which the polypeptide is expressed in the form of a proenzyme and the cell is cultured in the presence of a proteolytic enzyme capable of converting the proenzyme of the polypeptide into the mature polypeptide.

62. Use of a polypeptide according to any of claims 1-21 or 26-42 for reducing the allergenicity of a protein.

63. Use of a polypeptide according to any of claims 1-21 or 26-42 in food or feed manufacturing.

64. Use of a polypeptide according to any of claims 1-21 or 26-42 in the preparation of an enzyme preparation for use in food or feed manufacturing.

65. The use according to claim 63 for the reduction of allergens in food or feed, e.g. gluten or milk based products, including beverages, such as infant formula and dietary drinks.

66. Use according to claim 63 for increasing the digestibility of food or feed, such as, e.g., for increasing the digestibility of milk or wheat based food or feed products.

67. Use of a polypeptide according to any of claims 1-21 or 26-42 in the manufacturing of a cosmetic product.

68. Use of a polypeptide according to any of claims 1-21 or 26-42 treating scleroproteins, such as e.g. human or animal hair, cleaning fabrics.

69. A composition comprising a polypeptide as defined in any of the claims 1-21 or 26-42.

70. The composition according to claim 69 additionally comprising at least another enzyme, such as a protease, an amylase, a lipase, a hydrolase, a peroxidase, a cellulase, a transglutaminase, a glucose oxidase, a xylanase, a pectin methyl esterase or a pectin lyase.

71. A food additive comprising a polypeptide as defined in any of claims 1-21 or 26-42.

72. A cosmetic comprising a polypeptide as defined in any of claims 1-21 or 26-42.

Figure 1: Construction of pCaHj 527:

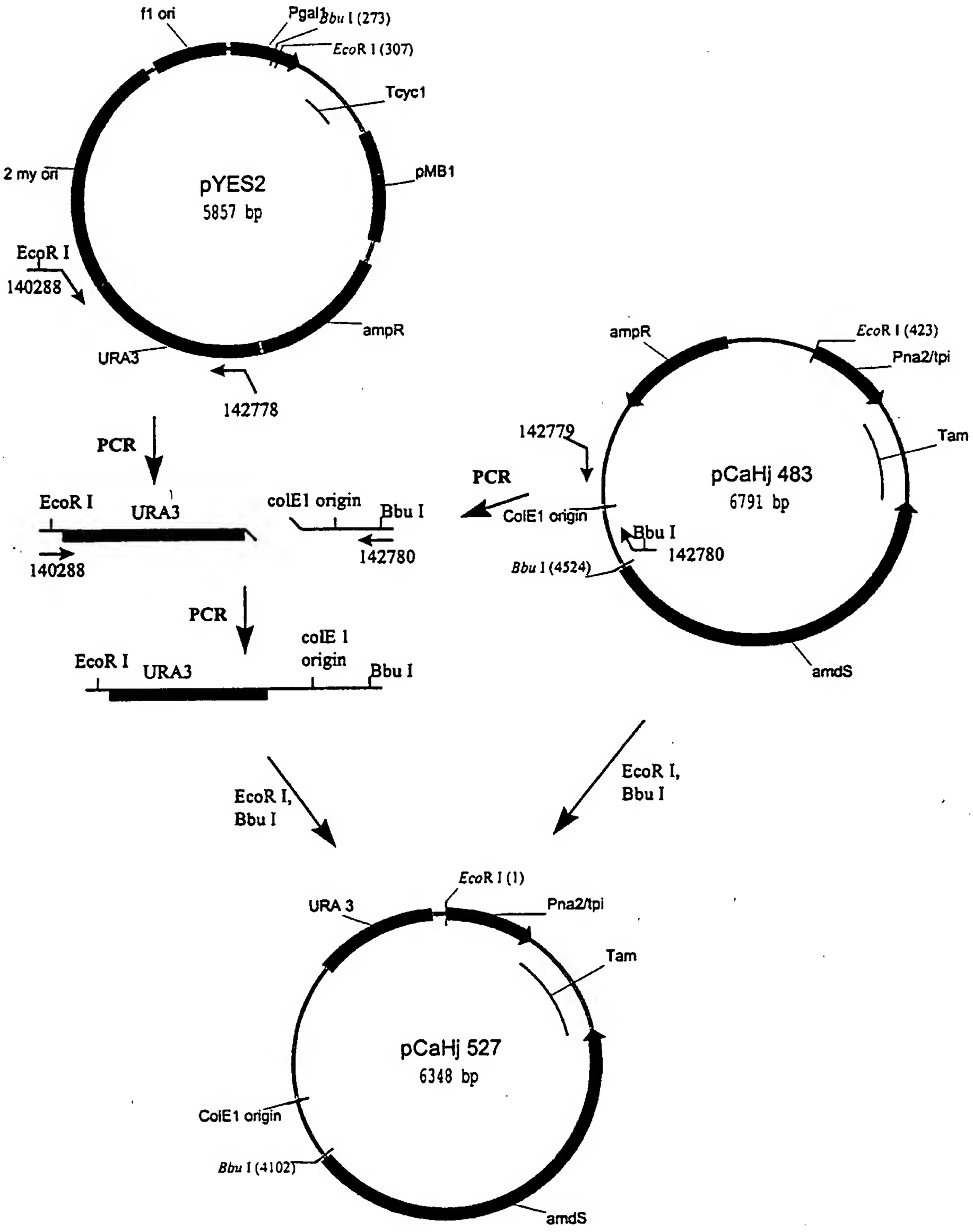


Figure 2: Construction of pCaHj 548

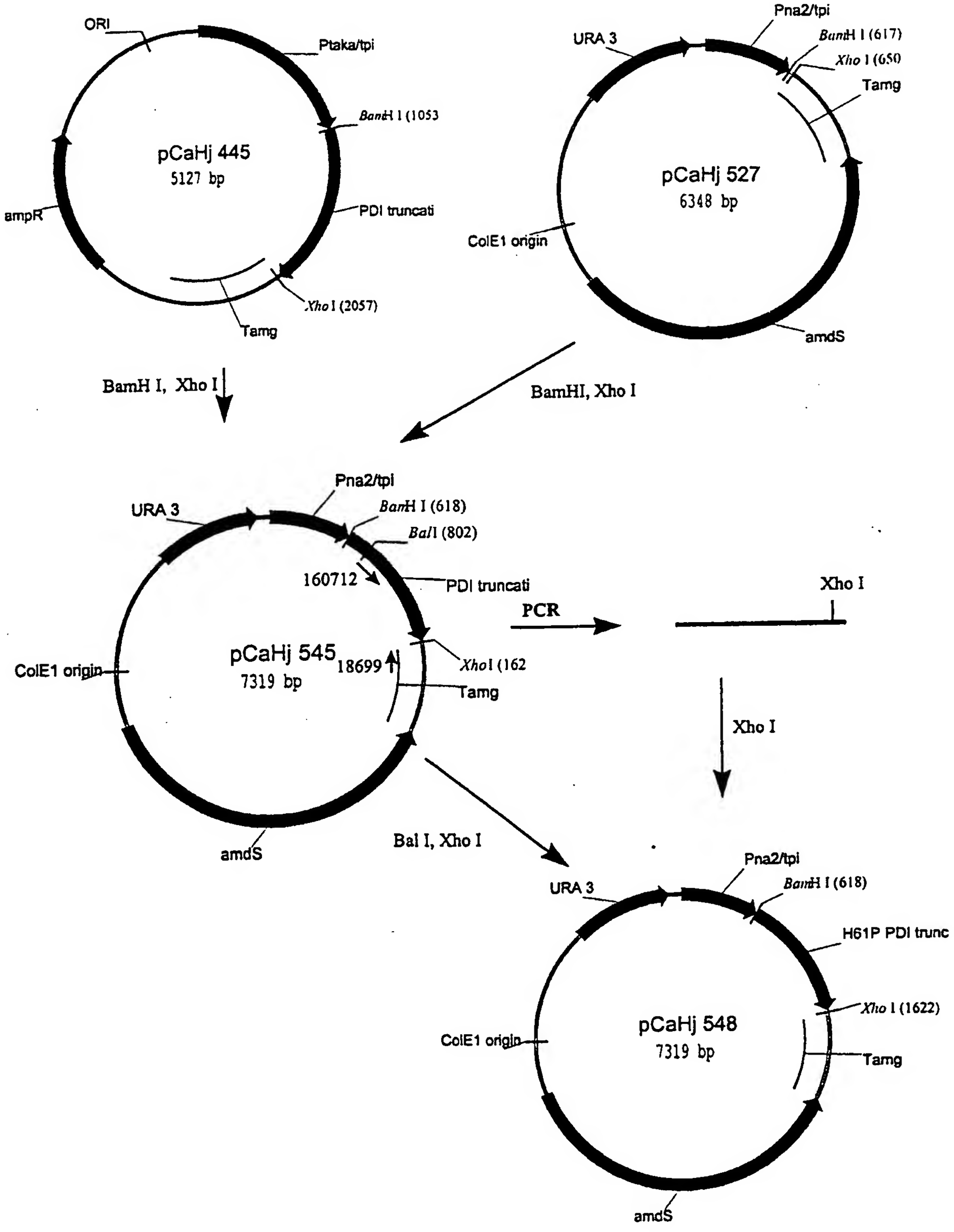
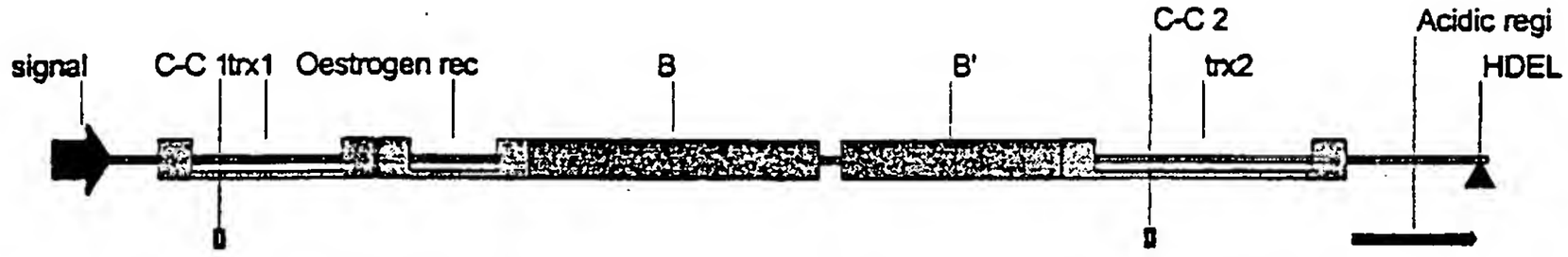


Figure 3



A oryzae
PD515

Figure 4

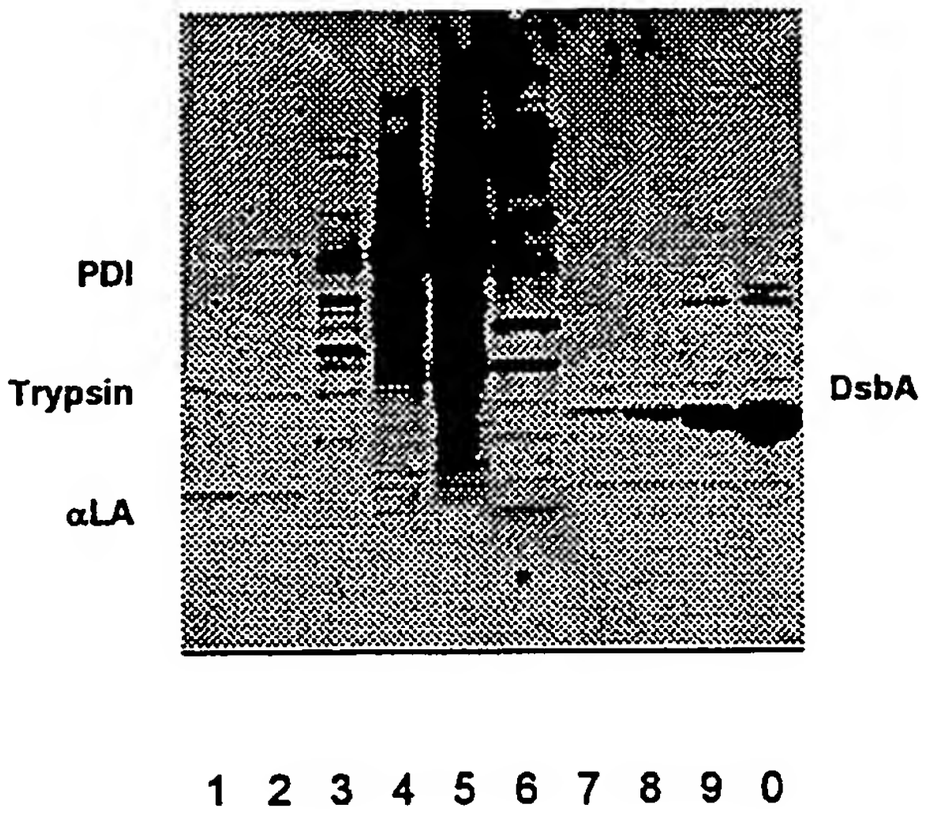
	T. reseii	*A. oryzae*	*H. insolens*	*S. cerevisiae*	*A. niger*
T. reseii	***	58.6	71.7	35.9	56.6
A. oryzae	55.7	***	59.8	33.6	77.9
H. insolens	34.1	54.4	***	36.6	58.4
S. cerevisiae	104.5	114.4	102.6	***	33.8
A. niger	58.7	23.8	58.6	112.2	***
	T. reseii	*A. oryzae*	*H. insolens*	*S. cerevisiae*	*A. niger*

```
                          1                                                  50
    A_ORYZAE.PRO    (1)   MRTFAPWILS---LLGASAVASAADATAEAPSDVVSLTGDTFETFVKEHD
    T_RESEII.PRO    (1)   MQQ-KRLTAA---LVAALAAVVSAE------SDVKSLTKDTFNDFINSND
  H_INSOLENS.PRO    (1)   MHKAQKFALG---LLAA--AAVATA------SDVVQLKKDTFDDFIKTND
S_CEREVISIAE.PRO    (1)   MKFSAGAVLSWSSLLLASSVFAQQEAVAPEDSAVVKLATDSFNEFIQSHD
    A_NIGER.PRO     (1)   MRSFAPWLVS---LLGASAVVAAADTE----SDVISLDQDTFESFMNEHG
      Consensus     (1)   MR  A WILS   LLAASAV AAAD      SDVVSL  DTFEDFINSHD
                          51                                                 100
    A_ORYZAE.PRO   (48)   LVLAEFFAPWCGHCKALAPKYEQAATELKEKNIPLVKVDCTEEEALCRDQ
    T_RESEII.PRO   (41)   LVLAESFAPWCGHCKALAPEYEEAATTLKDKSIKLAKVDCVEEADLCKEH
  H_INSOLENS.PRO   (40)   LVLAEFFAPWCGHCKALAPEYEEAATTLKEKNIKLAKVDCTEETDLCQQH
S_CEREVISIAE.PRO   (51)   LVLAEFFAPWCGHCKNMAPEYVKAAETLVEKNITLAQLDCTENQDLCMEH
    A_NIGER.PRO    (44)   LVLAEFFAPWCGHCKALAPKYEEAATELKAKNIPLVKVDCTAEEDLCRSQ
      Consensus    (51)   LVLAEFFAPWCGHCKALAPEYEEAATTLKEKNI LAKVDCTEE DLCREH
                          101                                                150
    A_ORYZAE.PRO   (98)   GVEGYPTLKIFR--GLDAVKPYQGARQTEAIVSYMVKQSLPAVSPVTPEN
    T_RESEII.PRO   (91)   GVEGYPTLKVFR--GLDKVAPYTGPRKADGITSYMVKQSLPAVSALTKDT
  H_INSOLENS.PRO   (90)   GVEGYPTLKVFR--GLDNVSPYKGQRKAAAITSYMIKQSLPAVSEVTKDN
S_CEREVISIAE.PRO  (101)   NIPGFPSLKIFKNSDVNNSIDYEGPRTAEAIVQFMIKQSQPAVAVVA--D
    A_NIGER.PRO    (94)   GVEGYPTLKIFR--GVDSSKPYQGARQTESIVSYMIKQSLPAVSSVNEEN
      Consensus   (101)   GVEGYPTLKIFR  GLD V PY G R AEAIVSYMIKQSLPAVS VT DN
                          151                                                200
    A_ORYZAE.PRO  (146)   LEEIKTMDKIVVIGYIASD--DQTANDIFTTFAESQRDNYLFAATSDAST
    T_RESEII.PRO  (139)   LEDFKTADKVVLVAYIAAD--DKASNETFTALANELRDTYLFGGVNDAAV
  H_INSOLENS.PRO  (138)   LEEFKKADKAVLVAYVDAS--DKASSFVFTQVAEKLRDNYPIGSSSDAAI
S_CEREVISIAE.PRO  (149)   LPAYLANETFVTPVIVQSGKIDADFNATFYSMANKHFNDYDFVSAENADD
    A_NIGER.PRO   (142)   LEEIKTMDKIVVIGYIPSD--DQETYQAFEKYAESQRDNYLFAATDDAAI
      Consensus   (151)   LEEFKT DKIVLIAYI SD  D  SNE FT LAE  RDNYLFAAT DAAI
                          201                                                250
    A_ORYZAE.PRO  (194)   AKAEGVKQFSIVLYKDFDEKKATYDGE----IEQDAILSWVKTASTPLVG
    T_RESEII.PRO  (187)   AEAEGVKFFSIVLYKSFDEGKNVFSEK----FDAEAIRNFAQVAATPLVG
  H_INSOLENS.PRO  (186)   AEAEGVKAPAIVLYKDFDEGKAVFSEK----FEVEAIEKFAKTGATPLIG
S_CEREVISIAE.PRO  (199)   DFKLSIYLPSAM------DEPVVYNGKKADIADIVFEKWLQVEALPYFG
    A_NIGER.PRO   (190)   AKSEGVEQFSIVLYKDFDEKKAVYDGE----IEQEAIHSWVKSASTPLVG
      Consensus   (201)   A AEGVK PSIVLYKDFDE KAVY GK     E EAI  WVKTAATPLVG
```

Figure 4 (continued)

```
                          251                                                300
  A_ORYZAE.PRO   (240) EIGPETYSGYIIAGIPLAYIFAETKEEREQYIEEFKFIAEKHKGSINIVT
  T_RESEII.PRO   (233) EVGPETYAGYMSAGIPLAYIFAETAEERENLAKTLKPVAEKYKGKINFAT
H_INSOLENS.PRO   (232) EIGPETYSDYMSAGIPLAYIFAETAEERKELSDKLKPIAEAQRGVINFGT
S_CEREVISIAE.PRO (243) EIDGSVIAQYVESGLPLGYLFYNDEEELEEYKPLFTELAKKNRGLMNFVS
  A_NIGER.PRO    (236) EIGPETYSGYIGAGVPLAYIFAETKEEREKYIEDFKPIAQKHKGAINIAT
    Consensus    (251) EIGPETYSGYISAGIPLAYIFAET EERE YTE FKPIAEKHKG INFAT
                          301                                                350
  A_ORYZAE.PRO   (290) IDAKLYGAHAGNLNLDPSKFPAFAIQDPEKNAKY--------PYDQSKE-
  T_RESEII.PRO   (283) IDAKNFGSHAGNINLKTDKFPAFALHDIEKNLKF--------PFDQSKE-
H_INSOLENS.PRO   (282) IDAKAFGAHAGNLNLKTDKFPAFAIQEVAKNQKF--------PFDQEKE-
S_CEREVISIAE.PRO (293) IDARKFGRHAGNLNMK-EQFPLFAIHDMTEDLKYGLPQLSEEAFDELSDK
  A_NIGER.PRO    (286) IDAKMFGAHAGNLNLDSQKFPAFAIQDPAKNAKY--------PYDQAKE-
    Consensus    (301) IDAK FGAHAGNLNLKTDKFPAFAIQDI KN KY         PFDQSKE
                          351                                                400
  A_ORYZAE.PRO   (331) --VKAKDIGKFIQDVLDDKVEPSIKSEAIPETQEGPVTVVVAHSYKDIVL
  T_RESEII.PRO   (324) --ITEKDIAAFVDGFSSGKIEASIKSEPIPETQEGPVTVVVAHSYKDIVL
H_INSOLENS.PRO   (323) --ITFEAIKAFVDDFVAGKIEPSIKSEPIPEKQEGPVTVVVAKNYNEIVL
S_CEREVISIAE.PRO (342) IVIESKAIESLVKDFLKGDASPIVKSQEIFENQDSSVFQLVGKNHDEIVN
  A_NIGER.PRO    (327) --INADEVEKFIQDVLDGKVEPSIKSEPVPESQEGPVTVVVAHSYKDIVI
    Consensus    (351)   I AKDI AFV DFL GKIEPSIKSEPIPETQEGPVTVVVAHSYKDIVL
                          401                                                450
  A_ORYZAE.PRO   (379) DNEKDVLIEFYAPWCGHCKALAPKYEELASLYKDIPI----VTIAKIDAT
  T_RESEII.PRO   (372) DDKKDVLIEFYTPWCGHCKALAPKYDELASLYAK-SDFKDKVVIAKVDAT
H_INSOLENS.PRO   (371) DDTKDVLIEFYAPWCGHCKALAPKYEELGALYAK-SEFKDRVVIAKVDAT
S_CEREVISIAE.PRO (392) DPKKDVLVLYYAPWCGHCKRLAPTYQELADTYANAT---SDVLIAKLDHT
  A_NIGER.PRO    (375) DNDKDVLIEFYAPWCGHCKALAPKYDELAALYADHPDLAAKVTIAKIDAT
    Consensus    (401) D  KDVLIEFYAPWCGHCKALAPKYDELAALYA  SD   KVVIAKIDAT
                          451                                                500
  A_ORYZAE.PRO   (425) ANDVPDS-ITGFPTIKLIAAGAKDSPVEYEGSRTVEDLANFVKENGKHKV
  T_RESEII.PRO   (421) ANDVPDE-IQGFPTIKLYPAGDKKNPVTYSGARTVEDFIEFIKENGKYKA
H_INSOLENS.PRO   (420) ANDVPDE-IQGFPTIKLYPAGAKGQPVTYSGSRTVEDLIKFIAENGKYKA
S_CEREVISIAE.PRO (439) ENDVRGVVIEGYPTIVLYPGGKKSESVVYQGSRSLDSLFDFIKENGHFDV
  A_NIGER.PRO    (425) ANDVPDP-ITGFPTLRLYPAGAKDSPIEYSGSRTVEDLANFVKENGKHNV
    Consensus    (451) ANDVPD  I GFPTIKLYPAGAK  PV YSGSRTVEDL  FIKENGKHKV
                          501                                          542
  A_ORYZAE.PRO   (474) DAIEVDPKKEQESGDATETRAASDETETPAATSDDKSEHDEL
  T_RESEII.PRO   (470) -GVETPAE----PTEEAEASESKASEEAKASEET----HDEL
H_INSOLENS.PRO   (469) -AISEDAEETSSATETTTETATKSEEAAKETATE----HDEL
S_CEREVISIAE.PRO (489) DGKAIYEEAQEKAAEEADADAELAD--------EEDAIHDEL
  A_NIGER.PRO    (474) DAINVASEETQEGGDVTEAAPSATEAETPAATDDEKAEHDEL
    Consensus    (501) DAL V AE    AGE TEA AS AE E  AA  DE A HDEL
```

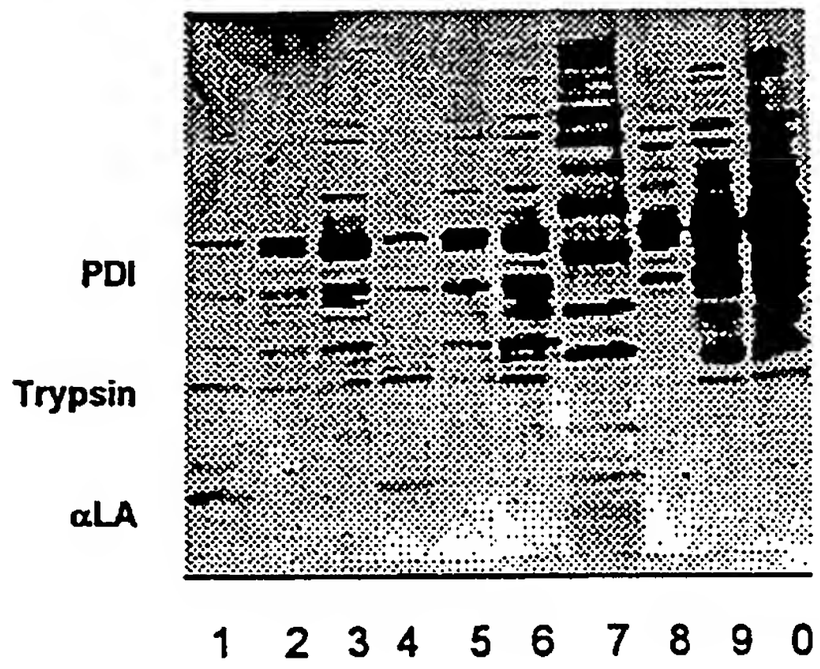
Figure 5



1. 5mg/ml α-lactalbumin (αLA) + trypsin
2. αLA + 0.04 mg/ml PDI + trypsin
3. αLA + 0.20 mg/ml PDI + trypsin
4. αLA + 1.00 mg/ml PDI + trypsin
5. αLA + 4.00 mg/ml PDI + trypsin
6. markers
7. αLA + 0.04 mg/ml DsbA + trypsin
8. αLA + 0.20 mg/ml DsbA + trypsin
9. αLA + 1.00 mg/ml DsbA + trypsin
0. αLA + 4.00 mg/ml DsbA + trypsin

Incubation was overnight at 37 °C.

Figure 6



1. 10 mg/ml αLA + 0.08 mg/ml PDI + trypsin
2. αLA + 0.20 mg/ml PDI + trypsin
3. αLA + 0.40 mg/ml PDI + trypsin
4. 20 mg/ml αLA + 0.16 mg/ml PDI + trypsin
5. αLA + 0.40 mg/ml PDI + trypsin
4. αLA + 0.80 mg/ml PDI + trypsin
5. markers
8. 0.08 mg/ml PDI
9. 0.40 mg/ml PDI
0. 0.80 mg/ml PDI

Incubation was overnight at 37 °C.

Figure 7

```
                1                                                    50
Pdi_Mouse  .......MLS RALLCLALAW AARVGADALE EEDNVLVLKK SNFEEALAAH
Pdi_Rat    .......MLS RALLCLALAW AARVGADALE EEDNVLVLKK SNFAEALAAH
Pdi_Bovin  .......MLR RALLCLALTA LFRAGAGAPD EEDHVLVLHK GNFDEALAAH
Pdi_Human  .......MLR RALLCLAVAA LVR..ADAPE EEDHVLVLRK SNFAEALAAH
Pdi_Rabit  .......MLR RAVLCLALAV TA.GWAWAAE EEDNVLVLKS SNFAEELAAH
Pdi_Chick  ..........  ..........  ......EPLE EEDGVLVLRA ANFEQALAAH
Pdi_Yeast  MKFSAGAVLS WSSLLLASSV FAQQEAVAPE DSA.VVKLAT DSFNEYIQSH

               51                                                   100
Pdi_Mouse  KYLLVEFYAP WCGHCKALAP EYAKRAAKLK AEGSEIRLAK VDATEESDLA
Pdi_Rat    NYLLVEFYAP WCGHCKALAP EYAKAAAKLK AEGSEIRLAK VDATEESDLA
Pdi_Bovin  KYLLVEFYAP WCGHCKALAP EYAKAAGKLK AEGSEIRLAK VDATEESDLA
Pdi_Human  KYLLVEFYAP WCGHCKALAP EYAKAAGKLK AEGSEIRLAK VDATEESDLA
Pdi_Rabit  KHLLVEFYAP WCGHCKALAP EYAKAAGKLK AEGSDIRLAK VDATEESDLA
Pdi_Chick  RHLLVEFYAP WCGHCKALAP EYAKAAAQLK AEGSEIRLAK VDATEEAELA
Pdi_Yeast  DLVLAEFFAP WCGHCKNMAP EYVKAAETL. .VEKNITLAQ IDCTENQDLC

              101                                                   150
Pdi_Mouse  QQYGVRGYPT IKFFKNGDTA SPKEYTAGRE ADDIVNWLKK RTGPAATTLS
Pdi_Rat    QQYGVRGYPT IKFFKNGDTA SPKEYTAGRE ADDIVNWLKK RTGPAATTLS
Pdi_Bovin  QQYGVRGYPT IKFFKNGDTA SPKEYTAGRE ADDIVNWLKK RTGPAASTLS
Pdi_Human  QQYGVRGYPT IKFFRNGDTA SPKEYTAGRE ADDIVNWLKK RTGPAATTLR
Pdi_Rabit  QQYGVRGYPT IKFFKNGDTA SPKEYTAGRE ADDIVNWLKK RTGPAATTLA
Pdi_Chick  QQFGVRGYPT IKFFRNGDKA APREYTAGRE ADDIVSWLKK RTGPAATTLT
Pdi_Yeast  MEHNIPGFPS LKIFKNSDVN NSIDYEGPRT AEAIVQFMIK QSQPAVAVVA

              151                                                   200
Pdi_Mouse  DTAAAESLVD SSEVTVIGFF KDVESDSAKQ FLLAAEAIDD IPFGITSNSG
Pdi_Rat    DTAAAESLVD SSEVTVIGFF KDAGSDSAKQ FLLAAEAVDD IPFGITSNSD
Pdi_Bovin  DGAAAEALVE SSEVAVIGFF KDMESDSAKQ FFLAAEVIDD IPFGITSNSD
Pdi_Human  DGAAAESLVE SSEVAVIGFF KDVESDSAKQ FLQAAEAIDD IPFGITSNSD
Pdi_Rabit  DSAAAESLVE SSEVAVIGFF KDVESDAAKQ FLLAAEATDD IPFGLTASSD
Pdi_Chick  DAAAAETLVD SSEVVVIGFF KDVTSDAAKE FLLAAESVDD IPFGISSSAD
Pdi_Yeast  DLPAYLANET FVTPVIVQSG KIDADFNATF YSMANKHFND YDFVSAENAD

              201                                                   250
Pdi_Mouse  VFSKYQLDKD GVVLFKKFDE GR..NNFEGE ITKEKLLD.F IKHNQLPLVI
Pdi_Rat    VFSKYQLDKD GVVLFKKFDE GR..NNFEGE ITKEKLLD.F IKHNQLPLVI
Pdi_Bovin  VFSKYQLDKD GVVLFKKFDE GR..NNFEGE VTKEKLLD.F IKHNQLPLVI
Pdi_Human  VFSKYQLDKD GVVLFKKFDE GR..NNFEGE VTKENLLD.F IKHNQLPLVI
Pdi_Rabit  VFSRYQVHQD GVVLFKKFDE GR..NNFEGE VTKEKLLD.F IKHNQLPLVI
Pdi_Chick  VFSKYQLSQD GVVLFKKFDE GR..NNFEGD LTKDNLLN.F IKSNQLPLVI
Pdi_Yeast  ..DDFKL... SIYLPSAMDE PVVYNGKKAD IADADVFEKW LQVEALPYFG

              251                                                   300
Pdi_Mouse  EFTEQTAPKI FGGEIKTHIL LFLPKSVSDY DGKLSSFKRA AEGF..KGKI
Pdi_Rat    EFTEQTAPKI FGGEIKTHIL LFLPKSVSDY DGKLSNFKKA AEGF..KGKI
Pdi_Bovin  EFTEQTAPKI FGGEIKTHIL LFLPKSVSDY EGKLSNFKKA AESF..KGKI
Pdi_Human  EFTEQTAPKI FGGEIKTHIL LFLPKSVSDY DGKLSNFKTA AESF..KGKI
Pdi_Rabit  EFTEQTAPKI FGGEIKTHIL LFLPRSAADH DGKLSGFKQA AEGF..KGKI
Pdi_Chick  EFTEQTAPKI FGGEIKTHIL LFLPKSVSDY EGKLDNFKTA AGNF..KGKI
Pdi_Yeast  EIDGSVFAQY VESGLPLGYL FY......ND EEELEEYKPL FTELAKKNRG
```

Figure 7 (continued)

```
            301                                                      350
Pdi_Mouse   LFIFIDSDHT DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Rat     LFIFIDSDHT DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Bovin   LFIFIDSDHT DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Human   LFIFIDSDHT DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Rabit   LFIFIDSDHA DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Chick   LFIFIDSDHS DNQRILEFFG LKKEECPAVR LITLEEEM.. .......TKY
Pdi_Yeast   LMNFVSIDAR KFGRHAGNLN M.KEQFPLFA IHDMTEDLKY GLPQLSEEAF

            351                                                      400
Pdi_Mouse   KPESDELTAE K..ITEFCHR FLEGKIKPHL MSQEVPEDWD KQPVKVLVGA
Pdi_Rat     KPESDELTAE K..ITQFCHH FLEGKIKPHL MSQELPEDWD KQPVKVLVGK
Pdi_Bovin   KPESDELTAE K..ITEFCHR FLEGKIKPHL MSQELPDDWD KQPVKVLVGK
Pdi_Human   KPESEELTAE R..ITEFCHR FLEGKIKPHL MSQERAGDWD KQPVKVPVGK
Pdi_Rabit   KPESDELTAE G..ITEFCQR FLEGKIKPHL MSQELPEDWD RQPVKVLVGK
Pdi_Chick   KPESDDLTAD K..IKEFCNK FLEGKIKPHL MSQDLPEDWD KQPVKVLVGK
Pdi_Yeast   DELSDKIVLE SKAIESLVKD FLKGDASPIV KSQEIFENQD S.SVFQLVGK

            401                                                      450
Pdi_Mouse   NFEEVAFDEK KNVFVEFYAP WCGHCKQLAP IWDKLGETY. KDHENIIIAK
Pdi_Rat     NFEEVAFDEK KNVFVEFYAP WCGHCKQLAP IWDKLGETY. KDHENIVIAK
Pdi_Bovin   NFEEVAFDEK KNVFVEFYAP WCGHCKQLAP IWDKLGETY. KDHENIVIAK
Pdi_Human   NFEDVAFDEK KNVFVEFYAP WCGHCKQLAP IWDKLGETY. KDHENIVIAK
Pdi_Rabit   NFEEVAFDEK KNVFVEFYAP WCGHCKQLAP IWDKLGETY. KEHQDIVIAK
Pdi_Chick   NFEEVAFDEN KNVFVEFYAP WCGHCKQLAP IWDKLGETY. RDHENIVIAK
Pdi_Yeast   NHDEIVNDPK KDVLVLYYAP WCGHCKRLAP TYQELADTYA NATSDVLIAK

            451                                                      500
Pdi_Mouse   MDSTANEVEA VKVHSFPTLK FFPASADRTV IDYNGERTLD GFKKFLESGG
Pdi_Rat     MDSTANEVEA VKVHSFPTLK FFPASADRTV IDYNGERTLD GFKKFLESGG
Pdi_Bovin   MDSTANEVEA VKVHSFPTLK FFPASADRTV IDYNGERTLD GFKKFLESGG
Pdi_Human   MDSTANEVEA VKVHSFPTLK FFPASADRTV IDYNGERTLD GFKKFLESGG
Pdi_Rabit   MDSTANEVEA VKVHSFPTLK FFPAGPGRTV IDYNGERTLD GFKKFLESGG
Pdi_Chick   MDSTANEVEA VKIHSFPTLK FFPAGSGRNV IDYNGERTLE GFKKFLESGG
Pdi_Yeast   LDHTENDVRG VVIEGYPTIV LYPGGKKSES VVYQGSRSLD SLFDFIKENG

            501                                                  538
Pdi_Mouse   QDGAGDDEDL .DLEE..ALE PDMEE..DDD QKAVKDEL
Pdi_Rat     QDGAGDNDDL .DLEE..ALE PDMEE..DDD QKAVKDEL
Pdi_Bovin   QDGAGDDDDL EDLEE..AEE PDLEE..DDD QKAVKDEL
Pdi_Human   QDGAGDDDDL EDLEE..AEE PDMEE..DDD QKAVKDEL
Pdi_Rabit   QDGAGDEDGL EDLEE..AEE PDLEE..DDD QKAVRDEL
Pdi_Chick   QDGAAADDDL EDLET..DEE TDLEEGDDDE QKIQKDEL
Pdi_Yeast   HFDVDGKALY EEAQEKAAEE ADADAELADE EDAIHDEL
```

SEQUENCE LISTING

<110> NOVO NORDISK A/S

<120> Novel Protein Disulfide Isomerase variants

<130> 5896A

<140>
<141>

<160> 24

<170> PatentIn Ver. 2.1

<210> 1
<211> 57
<212> PRT
<213> Aspergillus oryzae

<400> 1

Met	Arg	Thr	Phe	Ala	Pro	Trp	Ile	Leu	Ser	Leu	Leu	Gly	Ala	Ser	Ala
1				5				10					15		
Val	Ala	Ser	Ala	Ala	Asp	Ala	Thr	Ala	Glu	Ala	Pro	Ser	Asp	Val	Val
			20					25					30		
Ser	Leu	Thr	Gly	Asp	Thr	Phe	Glu	Thr	Phe	Val	Lys	Glu	His	Asp	Leu
		35					40					45			
Val	Leu	Ala	Glu	Phe	Phe	Ala	Pro	Trp							
	50					55									

<210> 2
<211> 454
<212> PRT
<213> Aspergillus oryzae

<400> 2

Lys	Ala	Leu	Ala	Pro	Lys	Tyr	Glu	Gln	Ala	Ala	Thr	Glu	Leu	Lys	Glu
1				5				10					15		
Lys	Asn	Ile	Pro	Leu	Val	Lys	Val	Asp	Cys	Thr	Glu	Glu	Glu	Ala	Leu
			20					25					30		
Cys	Arg	Asp	Gln	Gly	Val	Glu	Gly	Tyr	Pro	Thr	Leu	Lys	Ile	Phe	Arg

1

```
          35                  40                  45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
     50                  55                  60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
 65                  70                  75                  80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
                 85                  90                  95

Ile Gly Tyr Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr
            100                 105                 110

Thr Phe Ala Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser
        115                 120                 125

Asp Ala Ser Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val
    130                 135                 140

Leu Tyr Lys Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile
145                 150                 155                 160

Glu Gln Asp Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu
                165                 170                 175

Val Gly Glu Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly
            180                 185                 190

Ile Pro Leu Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln
        195                 200                 205

Phe Thr Glu Glu Phe Lys Phe Ile Ala Glu Lys His Lys Gly Ser Ile
    210                 215                 220

Asn Ile Val Thr Ile Asp Ala Lys Leu Tyr Gly Ala His Ala Gly Asn
225                 230                 235                 240

Leu Asn Leu Asp Pro Ser Lys Phe Pro Ala Phe Ala Ile Gln Asp Pro
                245                 250                 255

Glu Lys Asn Ala Lys Tyr Pro Tyr Asp Gln Ser Lys Glu Val Lys Ala
            260                 265                 270

Lys Asp Ile Gly Lys Phe Ile Gln Asp Val Leu Asp Asp Lys Val Glu
        275                 280                 285

Pro Ser Ile Lys Ser Glu Ala Ile Pro Glu Thr Gln Glu Gly Pro Val
```

```
    290                 295                 300

Thr Val Val Val Ala His Ser Tyr Lys Asp Leu Val Leu Asp Asn Glu
305                 310                 315                 320

Lys Asp Val Leu Leu Glu Phe Tyr Ala Pro Trp Cys Gly His Cys Lys
                325                 330                 335

Ala Leu Ala Pro Lys Tyr Glu Glu Leu Ala Ser Leu Tyr Lys Asp Ile
            340                 345                 350

Pro Glu Val Thr Ile Ala Lys Ile Asp Ala Thr Ala Asn Asp Val Pro
        355                 360                 365

Asp Ser Ile Thr Gly Phe Pro Thr Ile Lys Leu Phe Ala Ala Gly Ala
    370                 375                 380

Lys Asp Ser Pro Val Glu Tyr Glu Gly Ser Arg Thr Val Glu Asp Leu
385                 390                 395                 400

Ala Asn Phe Val Lys Glu Asn Gly Lys His Lys Val Asp Ala Leu Glu
                405                 410                 415

Val Asp Pro Lys Lys Glu Gln Glu Ser Gly Asp Ala Thr Glu Thr Arg
            420                 425                 430

Ala Ala Ser Asp Glu Thr Glu Thr Pro Ala Ala Thr Ser Asp Asp Lys
        435                 440                 445

Ser Glu His Asp Glu Leu
    450
```

<210> 3
<211> 54
<212> PRT
<213> Aspergillus oryzae

<400> 3

```
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
  1               5                  10                  15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
            20                  25                  30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
        35                  40                  45
```

Gly Leu Asp Ala Val Lys
50

<210> 4
<211> 70
<212> PRT
<213> Aspergillus oryzae

<400> 4
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
1 5 10 15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
20 25 30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
35 40 45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
50 55 60

Ile Val Ser Tyr Met Val
65 70

<210> 5
<211> 80
<212> PRT
<213> Aspergillus oryzae

<400> 5
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
1 5 10 15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
20 25 30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
35 40 45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
50 55 60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val

65 70 75 80

<210> 6
<211> 82
<212> PRT
<213> Aspergillus oryzae

<400> 6
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
1 5 10 15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
20 25 30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
35 40 45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
50 55 60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
65 70 75 80

Thr Pro

<210> 7
<211> 102
<212> PRT
<213> Aspergillus oryzae

<400> 7
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
1 5 10 15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
20 25 30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
35 40 45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
50 55 60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
65 70 75 80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
85 90 95

Ile Gly Tyr Ile Ala Ser
100

<210> 8
<211> 113
<212> PRT
<213> Aspergillus oryzae

<400> 8
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
1 5 10 15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
20 25 30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
35 40 45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
50 55 60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
65 70 75 80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
85 90 95

Ile Gly Tyr Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr
100 105 110

Thr

<210> 9
<211> 220

<212> PRT
<213> Aspergillus oryzae

<400> 9

```
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
  1               5                  10                  15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
             20                  25                  30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
         35                  40                  45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
     50                  55                  60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
 65                  70                  75                  80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
                 85                  90                  95

Ile Gly Tyr Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr
            100                 105                 110

Thr Phe Ala Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser
        115                 120                 125

Asp Ala Ser Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val
    130                 135                 140

Leu Tyr Lys Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile
145                 150                 155                 160

Glu Gln Asp Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu
                165                 170                 175

Val Gly Glu Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly
            180                 185                 190

Ile Pro Leu Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln
        195                 200                 205

Phe Thr Glu Glu Phe Lys Phe Ile Ala Glu Lys His
    210                 215                 220
```

<210> 10
<211> 450
<212> PRT
<213> Aspergillus oryzae

<400> 10

```
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
  1               5                  10                  15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
             20                  25                  30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
         35                  40                  45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
     50                  55                  60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
 65                  70                  75                  80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
                 85                  90                  95

Ile Gly Tyr Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr
            100                 105                 110

Thr Phe Ala Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser
        115                 120                 125

Asp Ala Ser Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val
    130                 135                 140

Leu Tyr Lys Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile
145                 150                 155                 160

Glu Gln Asp Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu
                165                 170                 175

Val Gly Glu Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly
            180                 185                 190

Ile Pro Leu Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln
        195                 200                 205

Phe Thr Glu Glu Phe Lys Phe Ile Ala Glu Lys His Lys Gly Ser Ile
    210                 215                 220
```

```
Asn Ile Val Thr Ile Asp Ala Lys Leu Tyr Gly Ala His Ala Gly Asn
225                 230                 235                 240

Leu Asn Leu Asp Pro Ser Lys Phe Pro Ala Phe Ala Ile Gln Asp Pro
                245                 250                 255

Glu Lys Asn Ala Lys Tyr Pro Tyr Asp Gln Ser Lys Glu Val Lys Ala
            260                 265                 270

Lys Asp Ile Gly Lys Phe Ile Gln Asp Val Leu Asp Asp Lys Val Glu
        275                 280                 285

Pro Ser Ile Lys Ser Glu Ala Ile Pro Glu Thr Gln Glu Gly Pro Val
    290                 295                 300

Thr Val Val Val Ala His Ser Tyr Lys Asp Leu Val Leu Asp Asn Glu
305                 310                 315                 320

Lys Asp Val Leu Leu Glu Phe Tyr Ala Pro Trp Cys Gly His Cys Lys
                325                 330                 335

Ala Leu Ala Pro Lys Tyr Glu Glu Leu Ala Ser Leu Tyr Lys Asp Ile
            340                 345                 350

Pro Glu Val Thr Ile Ala Lys Ile Asp Ala Thr Ala Asn Asp Val Pro
        355                 360                 365

Asp Ser Ile Thr Gly Phe Pro Thr Ile Lys Leu Phe Ala Ala Gly Ala
    370                 375                 380

Lys Asp Ser Pro Val Glu Tyr Glu Gly Ser Arg Thr Val Glu Asp Leu
385                 390                 395                 400

Ala Asn Phe Val Lys Glu Asn Gly Lys His Lys Val Asp Ala Leu Glu
                405                 410                 415

Val Asp Pro Lys Lys Glu Gln Glu Ser Gly Asp Ala Thr Glu Thr Arg
            420                 425                 430

Ala Ala Ser Asp Glu Thr Glu Thr Pro Ala Ala Thr Ser Asp Asp Lys
        435                 440                 445

Ser Glu
    450
```

<210> 11

<211> 331
<212> PRT
<213> Aspergillus oryzae

<400> 11

```
Lys Ala Leu Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu
  1               5                  10                  15

Lys Asn Ile Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu
             20                  25                  30

Cys Arg Asp Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg
         35                  40                  45

Gly Leu Asp Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala
     50                  55                  60

Ile Val Ser Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val
 65                  70                  75                  80

Thr Pro Glu Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val
                 85                  90                  95

Ile Gly Tyr Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr
            100                 105                 110

Thr Phe Ala Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser
        115                 120                 125

Asp Ala Ser Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val
    130                 135                 140

Leu Tyr Lys Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile
145                 150                 155                 160

Glu Gln Asp Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu
                165                 170                 175

Val Gly Glu Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly
            180                 185                 190

Ile Pro Leu Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln
        195                 200                 205

Phe Thr Glu Glu Phe Lys Phe Ile Ala Glu Lys His Lys Gly Ser Ile
    210                 215                 220

Asn Ile Val Thr Ile Asp Ala Lys Leu Tyr Gly Ala His Ala Gly Asn
```

225 230 235 240

Leu Asn Leu Asp Pro Ser Lys Phe Pro Ala Phe Ala Ile Gln Asp Pro
245 250 255

Glu Lys Asn Ala Lys Tyr Pro Tyr Asp Gln Ser Lys Glu Val Lys Ala
260 265 270

Lys Asp Ile Gly Lys Phe Ile Gln Asp Val Leu Asp Asp Lys Val Glu
275 280 285

Pro Ser Ile Lys Ser Glu Ala Ile Pro Glu Thr Gln Glu Gly Pro Val
290 295 300

Thr Val Val Val Ala His Ser Tyr Lys Asp Leu Val Leu Asp Asn Glu
305 310 315 320

Lys Asp Val Leu Leu Glu Phe Tyr Ala Pro Trp
325 330

<210> 12
<211> 1953
<212> DNA
<213> Aspergillus oryzae

<400> 12
ccctgctgtc cccatagaca gtacacacgt catcctttga tattgtcaca cttgacaaat 60
tcccgacacc atgcggactt tcgcaccttg gatcttgagc cttctagggg cttctgctgt 120
agcttctgct gccgatgcga ctgccgaagc tccctccgat gtggtctcgc tcaccgggga 180
cacattcgaa actttcgtca aggagcatga cctagttttg gccgagtttt ttgctccctg 240
gtgtggccat tgcaaggctc tcgctccgaa atacgagcag gccgccactg agttaaagga 300
aaagaacatt ccgctggtca aggttgattg caccgaggaa gaggctcttt gtagggacca 360
aggtgttgaa ggttacccca cgctgaagat tttccgtggc cttgacgctg ttaagcctta 420
tcagggagct cgtcagaccg aggcgtaagt gtcacctgtt tgttagcctt gctcaaataa 480
tattgaccgc tagtatcata ggattgtttc atacatggtc aagcagtcac tacctgctgt 540
gtcccctgtc accccagaaa acctcgaaga gatcaagact atggacaaga ttgtcgttat 600
tggttatatc gcgtctgacg accagactgc caatgatata ttcaccactt ttgccgagtc 660
acagagagac aactacctct tcgccgccac aagtgatgca tcgatcgcta aggcagaagg 720
tgttaagcaa ccttcgattg ttctctataa agacttcgat gaaaagaaag ctacttatga 780
tggagagatt gaacaggatg ccctcctcag ttgggtcaag actgccagta ccccccttggt 840
gggcgagctg ggcccagaga cttactccgg atatataacg gtatgtcaca agacacaatc 900
tcaatatcgc ttcacaacgt ttagtaaata atcatgagtt tctgacatgg gtttggttaa 960
ggctggcatt ccactggcgt acattttcgc cgaaaccaaa gaagagcgtg agcagttcac 1020
cgaggagttc aagttcatcg ccgagaaaca caagggttcc atcaatattg tcaccattga 1080
cgccaagttg tacggcgctc atgcaggcaa tctcaacctt gacccctcca agttccctgc 1140
attcgctatt caagaccctg aaaagaacgc caagtatcct tatgaccagt cgaaggaagt 1200

caaggccaag gatatcggta aattcatcca agacgttctt gatgataaag tagagccaag 1260
cattaagtct gaggctattc ctgagactca ggaaggtcct gttactgttg ttgtcgcgca 1320
ttcctataag gatctcgtcc ttgacaacga gaaggacgtc cttctcgaat tttatgcgcc 1380
atggtgcgga cactgcaagg cgtaagtttt ccccctcttt ctctacaacg aattatatcc 1440
actctcgctt gcgaatacct aattaaacct tgaatagcct tgccccgaag tacgaggaac 1500
ttgcaagcct ttacaaggat attcctgaag ttaccatcgc caaaattgac gcaacggcca 1560
acgatgtccc cgactccatt acaggatttc ctactattaa gctcttcgct gccggcgcca 1620
aggactcccc agttgaatat gaaggctctc gcacggtgga ggacctcgcc aacttcgtca 1680
aggagaatgg caagcacaag gtcgatgctc ttgaagttga tccgaagaaa gaacaggaga 1740
gtggcgatgc caccgagact cgggccgcct ctgacgagac cgaaactcct gctgctacta 1800
gcgatgacaa gtctgagcat gatgaattgt aaatttcatt tggcctgata gtttgatcca 1860
tatttatgtg aattcttgta ttctaccagc agtttgagca atcgcagcta cttccggctt 1920
aggaaactgt tgttctatcc tagtgggaag ctt 1953

<210> 13
<211> 515
<212> PRT
<213> Aspergillus oryzae

<400> 13

Met	Arg	Thr	Phe	Ala	Pro	Trp	Ile	Leu	Ser	Leu	Leu	Gly	Ala	Ser	Ala
1				5				10					15		
Val	Ala	Ser	Ala	Ala	Asp	Ala	Thr	Ala	Glu	Ala	Pro	Ser	Asp	Val	Val
			20					25					30		
Ser	Leu	Thr	Gly	Asp	Thr	Phe	Glu	Thr	Phe	Val	Lys	Glu	His	Asp	Leu
		35					40					45			
Val	Leu	Ala	Glu	Phe	Phe	Ala	Pro	Trp	Cys	Gly	His	Cys	Lys	Ala	Leu
	50					55					60				
Ala	Pro	Lys	Tyr	Glu	Gln	Ala	Ala	Thr	Glu	Leu	Lys	Glu	Lys	Asn	Ile
65					70					75					80
Pro	Leu	Val	Lys	Val	Asp	Cys	Thr	Glu	Glu	Glu	Ala	Leu	Cys	Arg	Asp
				85					90					95	
Gln	Gly	Val	Glu	Gly	Tyr	Pro	Thr	Leu	Lys	Ile	Phe	Arg	Gly	Leu	Asp
			100					105					110		
Ala	Val	Lys	Pro	Tyr	Gln	Gly	Ala	Arg	Gln	Thr	Glu	Ala	Ile	Val	Ser
		115					120					125			
Tyr	Met	Val	Lys	Gln	Ser	Leu	Pro	Ala	Val	Ser	Pro	Val	Thr	Pro	Glu
	130					135					140				

```
Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val Ile Gly Tyr
145                 150                 155                 160

Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr Thr Phe Ala
                165                 170                 175

Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser Asp Ala Ser
            180                 185                 190

Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val Leu Tyr Lys
        195                 200                 205

Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile Glu Gln Asp
    210                 215                 220

Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu Val Gly Glu
225                 230                 235                 240

Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly Ile Pro Leu
                245                 250                 255

Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln Phe Thr Glu
            260                 265                 270

Glu Phe Lys Phe Ile Ala Glu Lys His Lys Gly Ser Ile Asn Ile Val
        275                 280                 285

Thr Ile Asp Ala Lys Leu Tyr Gly Ala His Ala Gly Asn Leu Asn Leu
    290                 295                 300

Asp Pro Ser Lys Phe Pro Ala Phe Ala Ile Gln Asp Pro Glu Lys Asn
305                 310                 315                 320

Ala Lys Tyr Pro Tyr Asp Gln Ser Lys Glu Val Lys Ala Lys Asp Ile
                325                 330                 335

Gly Lys Phe Ile Gln Asp Val Leu Asp Asp Lys Val Glu Pro Ser Ile
            340                 345                 350

Lys Ser Glu Ala Ile Pro Glu Thr Gln Glu Gly Pro Val Thr Val Val
        355                 360                 365

Val Ala His Ser Tyr Lys Asp Leu Val Leu Asp Asn Glu Lys Asp Val
    370                 375                 380

Leu Leu Glu Phe Tyr Ala Pro Trp Cys Gly His Cys Lys Ala Leu Ala
385                 390                 395                 400
```

Pro Lys Tyr Glu Glu Leu Ala Ser Leu Tyr Lys Asp Ile Pro Glu Val
405 410 415

Thr Ile Ala Lys Ile Asp Ala Thr Ala Asn Asp Val Pro Asp Ser Ile
420 425 430

Thr Gly Phe Pro Thr Ile Lys Leu Phe Ala Ala Gly Ala Lys Asp Ser
435 440 445

Pro Val Glu Tyr Glu Gly Ser Arg Thr Val Glu Asp Leu Ala Asn Phe
450 455 460

Val Lys Glu Asn Gly Lys His Lys Val Asp Ala Leu Glu Val Asp Pro
465 470 475 480

Lys Lys Glu Gln Glu Ser Gly Asp Ala Thr Glu Thr Arg Ala Ala Ser
485 490 495

Asp Glu Thr Glu Thr Pro Ala Ala Thr Ser Asp Asp Lys Ser Glu His
500 505 510

Asp Glu Leu
515

<210> 14
<211> 843
<212> DNA
<213> Aspergillus oryzae

<400> 14
atgcggactt tcgcaccttg gatcttgagc cttctagggg cttctgctgt agcttctgct 60
gccgatgcga ctgccgaagc tccctccgat gtggtctcgc tcaccgggga cacattcgaa 120
actttcgtca aggagcatga cctagttttg gccgagtttt ttgctccctg gtgtggccct 180
tgcaaggctc tcgctccgaa atacgagcag gccgccactg agttaaagga aaagaacatt 240
ccgctggtca aggttgattg caccgaggaa gaggctcttt gtagggacca aggtgttgaa 300
ggttacccca cgctgaagat tttccgtggc cttgacgctg ttaagcctta tcagggagct 360
cgtcagaccg aggcgattgt ttcatacatg gtcaagcagt cactacctgc tgtgtcccct 420
gtcaccccag aaaacctcga agagatcaag actatggaca agattgtcgt tattggttat 480
atcgcgtctg acgaccagac tgccaatgat atattcacca cttttgccga gtcacagaga 540
gacaactacc tcttcgccgc cacaagtgat gcatcgatcg ctaaggcaga aggtgttaag 600
caaccttcga ttgttctcta taaagacttc gatgaaaaga aagctactta tgatggagag 660
attgaacagg atgccctcct cagttgggtc aagactgcca gtacccccctt ggtgggcgag 720
ctgggcccag agacttactc cggatatata acggctggca ttccactggc gtacattttc 780
gccgaaacca aagaagagcg tgagcagttc accgaggagt tcaagttcat cgccgagaaa 840
cac 843

<210> 15
<211> 281
<212> PRT
<213> Aspergillus oryzae

<400> 15

```
Met Arg Thr Phe Ala Pro Trp Ile Leu Ser Leu Leu Gly Ala Ser Ala
  1               5                  10                  15

Val Ala Ser Ala Ala Asp Ala Thr Ala Glu Ala Pro Ser Asp Val Val
             20                  25                  30

Ser Leu Thr Gly Asp Thr Phe Glu Thr Phe Val Lys Glu His Asp Leu
         35                  40                  45

Val Leu Ala Glu Phe Phe Ala Pro Trp Cys Gly Pro Cys Lys Ala Leu
     50                  55                  60

Ala Pro Lys Tyr Glu Gln Ala Ala Thr Glu Leu Lys Glu Lys Asn Ile
 65                  70                  75                  80

Pro Leu Val Lys Val Asp Cys Thr Glu Glu Glu Ala Leu Cys Arg Asp
                 85                  90                  95

Gln Gly Val Glu Gly Tyr Pro Thr Leu Lys Ile Phe Arg Gly Leu Asp
            100                 105                 110

Ala Val Lys Pro Tyr Gln Gly Ala Arg Gln Thr Glu Ala Ile Val Ser
        115                 120                 125

Tyr Met Val Lys Gln Ser Leu Pro Ala Val Ser Pro Val Thr Pro Glu
    130                 135                 140

Asn Leu Glu Glu Ile Lys Thr Met Asp Lys Ile Val Val Ile Gly Tyr
145                 150                 155                 160

Ile Ala Ser Asp Asp Gln Thr Ala Asn Asp Ile Phe Thr Thr Phe Ala
                165                 170                 175

Glu Ser Gln Arg Asp Asn Tyr Leu Phe Ala Ala Thr Ser Asp Ala Ser
            180                 185                 190

Ile Ala Lys Ala Glu Gly Val Lys Gln Pro Ser Ile Val Leu Tyr Lys
        195                 200                 205

Asp Phe Asp Glu Lys Lys Ala Thr Tyr Asp Gly Glu Ile Glu Gln Asp
    210                 215                 220
```

Ala Leu Leu Ser Trp Val Lys Thr Ala Ser Thr Pro Leu Val Gly Glu
225 230 235 240

Leu Gly Pro Glu Thr Tyr Ser Gly Tyr Ile Thr Ala Gly Ile Pro Leu
245 250 255

Ala Tyr Ile Phe Ala Glu Thr Lys Glu Glu Arg Glu Gln Phe Thr Glu
260 265 270

Glu Phe Lys Phe Ile Ala Glu Lys His
275 280

<210> 16
<211> 285
<212> DNA
<213> Aspergillus oryzae

<400> 16
gccgatgcga ctgccgaagc tccctccgat gtggtctcgc tcaccgggga cacattcgaa 60
actttcgtca aggagcatga cctagttttg gccgagtttt ttgctccctg gtgtggccat 120
tgcaaggctc tcgctccgaa atacgagcag gccgccactg agttaaagga aaagaacatt 180
ccgctggtca aggttgattg caccgaggaa gaggctcttt gtagggacca aggtgttgaa 240
ggttacccca cgctgaagat tttccgtggc cttgacgctg ttaag 285

<210> 17
<211> 95
<212> PRT
<213> Aspergillus oryzae

<400> 17
Ala Asp Ala Thr Ala Glu Ala Pro Ser Asp Val Val Ser Leu Thr Gly
1 5 10 15

Asp Thr Phe Glu Thr Phe Val Lys Glu His Asp Leu Val Leu Ala Glu
20 25 30

Phe Phe Ala Pro Trp Cys Gly Pro Cys Lys Ala Leu Ala Pro Lys Tyr
35 40 45

Glu Gln Ala Ala Thr Glu Leu Lys Glu Lys Asn Ile Pro Leu Val Lys
50 55 60

Val Asp Cys Thr Glu Glu Glu Ala Leu Cys Arg Asp Gln Gly Val Glu
65 70 75 80

Gly Tyr Pro Thr Leu Lys Ile Phe Arg Gly Leu Asp Ala Val Lys
85 90 95

<210> 18
<211> 102
<212> PRT
<213> Aspergillus oryzae

<400> 18
Gly Pro Val Thr Val Val Val Ala His Ser Tyr Lys Asp Leu Val Leu
1 5 10 15

Asp Asn Glu Lys Asp Val Leu Leu Glu Phe Tyr Ala Pro Trp Cys Gly
20 25 30

His Cys Lys Ala Leu Ala Pro Lys Tyr Glu Glu Leu Ala Ser Leu Tyr
35 40 45

Lys Asp Ile Pro Glu Val Thr Ile Ala Lys Ile Asp Ala Thr Ala Asn
50 55 60

Asp Val Pro Asp Ser Ile Thr Gly Phe Pro Thr Ile Lys Leu Phe Ala
65 70 75 80

Ala Gly Ala Lys Asp Ser Pro Val Glu Tyr Glu Gly Ser Arg Thr Val
85 90 95

Glu Asp Leu Ala Asn Phe
100

<210> 19
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:142779

<400> 19
ttgaattgaa aatagattga tttaaaactt c 31

<210> 20

<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:142780

<400> 20
ttgcatgcgt aatcatggtc atagc 25

<210> 21
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:140288

<400> 21
ttgaattcat gggtaataac tgatat 26

<210> 22
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:142778

<400> 22
aaatcaatct attttcaatt caattcatca tt 32

<210> 23
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:160712

<400> 23
cccttgcaag gctctcgctc c 21

<210> 24

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:18699

<400> 24
ttgccctcat ccccatcctt t 21

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/DK 00/00265

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/61 C12N15/62 C12N9/90 A61K7/48 A61K7/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
STRAND, EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 00636 A (NOVONORDISK AS ;HJORT CARSTEN M (DK)) 5 January 1995 (1995-01-05) cited in the application page 5, line 31 -page 6, line 27 * SEQ ID No.3 * claims 30,AND,36 ---	1-61, 68-70,72
A	WO 95 01425 A (AASLYNG DORRIT ;HJORT CARSTEN M (DK); NOVONORDISK AS (DK); ANDERSE) 12 January 1995 (1995-01-12) cited in the application page 9, line 10 -page 10, line 35 ---	1-61, 68-70,72

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 October 2000	07/11/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mata Vicente, T.

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int ional Application No
PCT/DK 00/00265

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	LUNDSTRÖM, J. ET AL.: "A Pro to His Mutation in Active Site of Thioredoxin Increases Its Disulfide-Isomerase Activity 10-fold." J. BIOL. CHEM., vol. 267, no. 13, 5 May 1992 (1992-05-05), pages 9047-9052, XP002133750 abstract ----	1-61
A	US 5 792 506 A (LOZANO ROSA ET AL) 11 August 1998 (1998-08-11) abstract claim 3 -----	62-66, 69,71

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/DK 00/00265

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9500636 A	05-01-1995	AU 7068494 A	17-01-1995
		EP 0708825 A	01-05-1996
		JP 9500014 T	07-01-1997
		US 5879664 A	09-03-1999
WO 9501425 A	12-01-1995	AU 7068294 A	24-01-1995
		AU 7068394 A	24-01-1995
		WO 9501420 A	12-01-1995
		EP 0710281 A	08-05-1996
		JP 8511944 T	17-12-1996
US 5792506 A	11-08-1998	US 6113951 A	05-09-2000
		AU 3956395 A	15-05-1996
		EP 0784676 A	23-07-1997
		JP 10510244 T	06-10-1998
		WO 9612799 A	02-05-1996
		US 5952034 A	14-09-1999
		AT 175312 T	15-01-1999
		AU 677771 B	08-05-1997
		AU 2861792 A	21-05-1993
		CA 2121137 A	29-04-1993
		CZ 9400832 A	16-08-1995
		DE 69228130 D	18-02-1999
		EP 0672127 A	20-09-1995
		HU 69780 A	28-09-1995
		JP 7502887 T	30-03-1995
		NZ 244695 A	26-03-1996
		SK 41894 A	11-07-1995
		WO 9308274 A	29-04-1993
		US 6114504 A	05-09-2000
		EP 0863154 A	09-09-1998
		ZA 9207831 A	27-04-1993